LNCS 6769

Aaron Marcus (Ed.)

Design, User Experience, and Usability

Theory, Methods, Tools and Practice

First International Conference, DUXU 2011
Held as Part of HCI International 2011
Orlando, FL, USA, July 2011, Proceedings, Part I

Part I

HCI2011
INTERNATIONAL

Springer

Lecture Notes in Computer Science 6769

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK
Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA
Josef Kittler
University of Surrey, Guildford, UK
Jon M. Kleinberg
Cornell University, Ithaca, NY, USA
Alfred Kobsa
University of California, Irvine, CA, USA
Friedemann Mattern
ETH Zurich, Switzerland
John C. Mitchell
Stanford University, CA, USA
Moni Naor
Weizmann Institute of Science, Rehovot, Israel
Oscar Nierstrasz
University of Bern, Switzerland
C. Pandu Rangan
Indian Institute of Technology, Madras, India
Bernhard Steffen
TU Dortmund University, Germany
Madhu Sudan
Microsoft Research, Cambridge, MA, USA
Demetri Terzopoulos
University of California, Los Angeles, CA, USA
Doug Tygar
University of California, Berkeley, CA, USA
Gerhard Weikum
Max Planck Institute for Informatics, Saarbruecken, Germany

Aaron Marcus (Ed.)

Design, User Experience and Usability

Theory, Methods, Tools and Practice

First International Conference, DUXU 2011
Held as Part of HCI International 2011
Orlando, FL, USA, July 9-14, 2011
Proceedings, Part I

Springer

Volume Editor

Aaron Marcus
Aaron Marcus and Associates, Inc.
1196 Euclid Avenue, Suite 1F
Berkeley, CA, 94708-1640, USA
E-mail: Aaron.Marcus@AMandA.com

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-21674-9 e-ISBN 978-3-642-21675-6
DOI 10.1007/978-3-642-21675-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011929078

CR Subject Classification (1998): H.5, H.3-4, K.6, D.2, C.2, K.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

© Springer-Verlag Berlin Heidelberg 2011
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The 14th International Conference on Human–Computer Interaction, HCI International 2011, was held in Orlando, Florida, USA, July 9–14, 2011, jointly with the Symposium on Human Interface (Japan) 2011, the 9th International Conference on Engineering Psychology and Cognitive Ergonomics, the 6th International Conference on Universal Access in Human–Computer Interaction, the 4th International Conference on Virtual and Mixed Reality, the 4th International Conference on Internationalization, Design and Global Development, the 4th International Conference on Online Communities and Social Computing, the 6th International Conference on Augmented Cognition, the Third International Conference on Digital Human Modeling, the Second International Conference on Human-Centered Design, and the First International Conference on Design, User Experience, and Usability.

A total of 4,039 individuals from academia, research institutes, industry and governmental agencies from 67 countries submitted contributions, and 1,318 papers that were judged to be of high scientific quality were included in the program. These papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. The papers accepted for presentation thoroughly cover the entire field of human–computer interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas.

This volume, edited by Aaron Marcus, contains papers in the thematic area of design, user experience, and usability (DUXU), addressing the following major topics:

- DUXU theory, methods and tools
- DUXU guidelines and standards
- Novel DUXU: devices and their user interfaces
- DUXU in industry
- DUXU in the mobile and vehicle context

The remaining volumes of the HCI International 2011 Proceedings are:

- Volume 1, LNCS 6761, Human–Computer Interaction—Design and Development Approaches (Part I), edited by Julie A. Jacko
- Volume 2, LNCS 6762, Human–Computer Interaction—Interaction Techniques and Environments (Part II), edited by Julie A. Jacko
- Volume 3, LNCS 6763, Human–Computer Interaction—Towards Mobile and Intelligent Interaction Environments (Part III), edited by Julie A. Jacko
- Volume 4, LNCS 6764, Human–Computer Interaction—Users and Applications (Part IV), edited by Julie A. Jacko
- Volume 5, LNCS 6765, Universal Access in Human–Computer Interaction—Design for All and eInclusion (Part I), edited by Constantine Stephanidis

- Volume 6, LNCS 6766, Universal Access in Human–Computer Interaction—Users Diversity (Part II), edited by Constantine Stephanidis
- Volume 7, LNCS 6767, Universal Access in Human–Computer Interaction—Context Diversity (Part III), edited by Constantine Stephanidis
- Volume 8, LNCS 6768, Universal Access in Human–Computer Interaction—Applications and Services (Part IV), edited by Constantine Stephanidis
- Volume 10, LNCS 6770, Design, User Experience, and Usability—Understanding the User Experience (Part II), edited by Aaron Marcus
- Volume 11, LNCS 6771, Human Interface and the Management of Information—Design and Interaction (Part I), edited by Michael J. Smith and Gavriel Salvendy
- Volume 12, LNCS 6772, Human Interface and the Management of Information—Interacting with Information (Part II), edited by Gavriel Salvendy and Michael J. Smith
- Volume 13, LNCS 6773, Virtual and Mixed Reality—New Trends (Part I), edited by Randall Shumaker
- Volume 14, LNCS 6774, Virtual and Mixed Reality—Systems and Applications (Part II), edited by Randall Shumaker
- Volume 15, LNCS 6775, Internationalization, Design and Global Development, edited by P.L. Patrick Rau
- Volume 16, LNCS 6776, Human-Centered Design, edited by Masaaki Kurosu
- Volume 17, LNCS 6777, Digital Human Modeling, edited by Vincent G. Duffy
- Volume 18, LNCS 6778, Online Communities and Social Computing, edited by A. Ant Ozok and Panayiotis Zaphiris
- Volume 19, LNCS 6779, Ergonomics and Health Aspects of Work with Computers, edited by Michelle M. Robertson
- Volume 20, LNAI 6780, Foundations of Augmented Cognition: Directing the Future of Adaptive Systems, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
- Volume 21, LNAI 6781, Engineering Psychology and Cognitive Ergonomics, edited by Don Harris
- Volume 22, CCIS 173, HCI International 2011 Posters Proceedings (Part I), edited by Constantine Stephanidis
- Volume 23, CCIS 174, HCI International 2011 Posters Proceedings (Part II), edited by Constantine Stephanidis

I would like to thank the Program Chairs and the members of the Program Boards of all Thematic Areas, listed herein, for their contribution to the highest scientific quality and the overall success of the HCI International 2011 Conference.

In addition to the members of the Program Boards, I also wish to thank the following volunteer external reviewers: Roman Vilimek from Germany, Ramalingam Ponnusamy from India, Si Jung "Jun" Kim from the USA, and Ilia Adami, Iosif Klironomos, Vassilis Kouroumalis, George Margetis, and Stavroula Ntoa from Greece.

This conference would not have been possible without the continuous support and advice of the Conference Scientific Advisor, Gavriel Salvendy, as well as the dedicated work and outstanding efforts of the Communications and Exhibition Chair and Editor of HCI International News, Abbas Moallem.

I would also like to thank for their contribution toward the organization of the HCI International 2011 Conference the members of the Human–Computer Interaction Laboratory of ICS-FORTH, and in particular Margherita Antona, George Paparoulis, Maria Pitsoulaki, Stavroula Ntoa, Maria Bouhli and George Kapnas.

July 2011 Constantine Stephanidis

Organization

Ergonomics and Health Aspects of Work with Computers

Program Chair: Michelle M. Robertson

Arne Aarås, Norway
Pascale Carayon, USA
Jason Devereux, UK
Wolfgang Friesdorf, Germany
Martin Helander, Singapore
Ed Israelski, USA
Ben-Tzion Karsh, USA
Waldemar Karwowski, USA
Peter Kern, Germany
Danuta Koradecka, Poland
Nancy Larson, USA
Kari Lindström, Finland
Brenda Lobb, New Zealand
Holger Luczak, Germany
William S. Marras, USA
Aura C. Matias, Philippines
Matthias Rötting, Germany
Michelle L. Rogers, USA
Dominique L. Scapin, France
Lawrence M. Schleifer, USA
Michael J. Smith, USA
Naomi Swanson, USA
Peter Vink, The Netherlands
John Wilson, UK

Human Interface and the Management of Information

Program Chair: Michael J. Smith

Hans-Jörg Bullinger, Germany
Alan Chan, Hong Kong
Shin'ichi Fukuzumi, Japan
Jon R. Gunderson, USA
Michitaka Hirose, Japan
Jhilmil Jain, USA
Yasufumi Kume, Japan
Mark Lehto, USA
Hirohiko Mori, Japan
Fiona Fui-Hoon Nah, USA
Shogo Nishida, Japan
Robert Proctor, USA
Youngho Rhee, Korea
Anxo Cereijo Roibás, UK
Katsunori Shimohara, Japan
Dieter Spath, Germany
Tsutomu Tabe, Japan
Alvaro D. Taveira, USA
Kim-Phuong L. Vu, USA
Tomio Watanabe, Japan
Sakae Yamamoto, Japan
Hidekazu Yoshikawa, Japan
Li Zheng, P.R. China

Human–Computer Interaction

Program Chair: Julie A. Jacko

Sebastiano Bagnara, Italy
Sherry Y. Chen, UK
Marvin J. Dainoff, USA
Jianming Dong, USA
John Eklund, Australia
Xiaowen Fang, USA
Ayse Gurses, USA
Vicki L. Hanson, UK
Sheue-Ling Hwang, Taiwan
Wonil Hwang, Korea
Yong Gu Ji, Korea
Steven A. Landry, USA
Gitte Lindgaard, Canada
Chen Ling, USA
Yan Liu, USA
Chang S. Nam, USA
Celestine A. Ntuen, USA
Philippe Palanque, France
P.L. Patrick Rau, P.R. China
Ling Rothrock, USA
Guangfeng Song, USA
Steffen Staab, Germany
Wan Chul Yoon, Korea
Wenli Zhu, P.R. China

Engineering Psychology and Cognitive Ergonomics

Program Chair: Don Harris

Guy A. Boy, USA
Pietro Carlo Cacciabue, Italy
John Huddlestone, UK
Kenji Itoh, Japan
Hung-Sying Jing, Taiwan
Wen-Chin Li, Taiwan
James T. Luxhøj, USA
Nicolas Marmaras, Greece
Sundaram Narayanan, USA
Mark A. Neerincx, The Netherlands
Jan M. Noyes, UK
Kjell Ohlsson, Sweden
Axel Schulte, Germany
Sarah C. Sharples, UK
Neville A. Stanton, UK
Xianghong Sun, P.R. China
Andrew Thatcher, South Africa
Matthew J.W. Thomas, Australia
Mark Young, UK
Rolf Zon, The Netherlands

Universal Access in Human–Computer Interaction

Program Chair: Constantine Stephanidis

Julio Abascal, Spain
Ray Adams, UK
Elisabeth André, Germany
Margherita Antona, Greece
Chieko Asakawa, Japan
Christian Bühler, Germany
Jerzy Charytonowicz, Poland
Pier Luigi Emiliani, Italy
Michael Fairhurst, UK
Dimitris Grammenos, Greece
Andreas Holzinger, Austria
Simeon Keates, Denmark
Georgios Kouroupetroglou, Greece
Sri Kurniawan, USA
Patrick M. Langdon, UK
Seongil Lee, Korea

Zhengjie Liu, P.R. China
Klaus Miesenberger, Austria
Helen Petrie, UK
Michael Pieper, Germany
Anthony Savidis, Greece
Andrew Sears, USA
Christian Stary, Austria
Hirotada Ueda, Japan
Jean Vanderdonckt, Belgium
Gregg C. Vanderheiden, USA
Gerhard Weber, Germany
Harald Weber, Germany
Panayiotis Zaphiris, Cyprus

Virtual and Mixed Reality

Program Chair: Randall Shumaker

Pat Banerjee, USA
Mark Billinghurst, New Zealand
Charles E. Hughes, USA
Simon Julier, UK
David Kaber, USA
Hirokazu Kato, Japan
Robert S. Kennedy, USA
Young J. Kim, Korea
Ben Lawson, USA
Gordon McK Mair, UK
David Pratt, UK
Albert “Skip” Rizzo, USA
Lawrence Rosenblum, USA
Jose San Martin, Spain
Dieter Schmalstieg, Austria
Dylan Schmorrow, USA
Kay Stanney, USA
Janet Weisenford, USA
Mark Wiederhold, USA

Internationalization, Design and Global Development

Program Chair: P.L. Patrick Rau

Michael L. Best, USA
Alan Chan, Hong Kong
Lin-Lin Chen, Taiwan
Andy M. Dearden, UK
Susan M. Dray, USA
Henry Been-Lirn Duh, Singapore
Vanessa Evers, The Netherlands
Paul Fu, USA
Emilie Gould, USA
Sung H. Han, Korea
Veikko Ikonen, Finland
Toshikazu Kato, Japan
Esin Kiris, USA
Apala Lahiri Chavan, India
James R. Lewis, USA
James J.W. Lin, USA
Rungtai Lin, Taiwan
Zhengjie Liu, P.R. China
Aaron Marcus, USA
Allen E. Milewski, USA
Katsuhiko Ogawa, Japan
Oguzhan Ozcan, Turkey
Girish Prabhu, India
Kerstin Röse, Germany
Supriya Singh, Australia
Alvin W. Yeo, Malaysia
Hsiu-Ping Yueh, Taiwan

Online Communities and Social Computing

Program Chairs: A. Ant Ozok, Panayiotis Zaphiris

Chadia N. Abras, USA
Chee Siang Ang, UK
Peter Day, UK
Fiorella De Cindio, Italy
Heidi Feng, USA
Anita Komlodi, USA
Piet A.M. Kommers, The Netherlands
Andrew Laghos, Cyprus
Stefanie Lindstaedt, Austria
Gabriele Meiselwitz, USA
Hideyuki Nakanishi, Japan
Anthony F. Norcio, USA
Ulrike Pfeil, UK
Elaine M. Raybourn, USA
Douglas Schuler, USA
Gilson Schwartz, Brazil
Laura Slaughter, Norway
Sergei Stafeev, Russia
Asimina Vasalou, UK
June Wei, USA
Haibin Zhu, Canada

Augmented Cognition

Program Chairs: Dylan D. Schmorrow, Cali M. Fidopiastis

Monique Beaudoin, USA
Chris Berka, USA
Joseph Cohn, USA
Martha E. Crosby, USA
Julie Drexler, USA
Ivy Estabrooke, USA
Chris Forsythe, USA
Wai Tat Fu, USA
Marc Grootjen, The Netherlands
Jefferson Grubb, USA
Santosh Mathan, USA
Rob Matthews, Australia
Dennis McBride, USA
Eric Muth, USA
Mark A. Neerincx, The Netherlands
Denise Nicholson, USA
Banu Onaral, USA
Kay Stanney, USA
Roy Stripling, USA
Rob Taylor, UK
Karl van Orden, USA

Digital Human Modeling

Program Chair: Vincent G. Duffy

Karim Abdel-Malek, USA
Giuseppe Andreoni, Italy
Thomas J. Armstrong, USA
Norman I. Badler, USA
Fethi Calisir, Turkey
Daniel Carruth, USA
Keith Case, UK
Julie Charland, Canada
Yaobin Chen, USA
Kathryn Cormican, Ireland
Daniel A. DeLaurentis, USA
Yingzi Du, USA
Okan Ersoy, USA
Enda Fallon, Ireland
Yan Fu, P.R. China
Afzal Godil, USA

Ravindra Goonetilleke, Hong Kong
Anand Gramopadhye, USA
Lars Hanson, Sweden
Pheng Ann Heng, Hong Kong
Bo Hoege, Germany
Hongwei Hsiao, USA
Tianzi Jiang, P.R. China
Nan Kong, USA
Steven A. Landry, USA
Kang Li, USA
Zhizhong Li, P.R. China
Tim Marler, USA
Ahmet F. Ozok, Turkey
Srinivas Peeta, USA
Sudhakar Rajulu, USA
Matthias Rötting, Germany
Matthew Reed, USA
Johan Stahre, Sweden
Mao-Jiun Wang, Taiwan
Xuguang Wang, France
Jingzhou (James) Yang, USA
Gulcin Yucel, Turkey
Tingshao Zhu, P.R. China

Human-Centered Design

Program Chair: Masaaki Kurosu

Julio Abascal, Spain
Simone Barbosa, Brazil
Tomas Berns, Sweden
Nigel Bevan, UK
Torkil Clemmensen, Denmark
Susan M. Dray, USA
Vanessa Evers, The Netherlands
Xiaolan Fu, P.R. China
Yasuhiro Horibe, Japan
Jason Huang, P.R. China
Minna Isomursu, Finland
Timo Jokela, Finland
Mitsuhiko Karashima, Japan
Tadashi Kobayashi, Japan
Seongil Lee, Korea
Kee Yong Lim, Singapore
Zhengjie Liu, P.R. China
Loïc Martínez-Normand, Spain
Monique Noirhomme-Fraiture, Belgium
Philippe Palanque, France
Annelise Mark Pejtersen, Denmark
Kerstin Röse, Germany
Dominique L. Scapin, France
Haruhiko Urokohara, Japan
Gerrit C. van der Veer, The Netherlands
Janet Wesson, South Africa
Toshiki Yamaoka, Japan
Kazuhiko Yamazaki, Japan
Silvia Zimmermann, Switzerland

Design, User Experience, and Usability

Program Chair: Aaron Marcus

Ronald Baecker, Canada
Barbara Ballard, USA
Konrad Baumann, Austria
Arne Berger, Germany
Randolph Bias, USA
Jamie Blustein, Canada
Ana Boa-Ventura, USA
Lorenzo Cantoni, Switzerland
Sameer Chavan, Korea
Wei Ding, USA
Maximilian Eibl, Germany
Zelda Harrison, USA

Rüdiger Heimgärtner, Germany
Brigitte Herrmann, Germany
Sabine Kabel-Eckes, USA
Kaleem Khan, Canada
Jonathan Kies, USA
Jon Kolko, USA
Helga Letowt-Vorbek, South Africa
James Lin, USA
Frazer McKimm, Ireland
Michael Renner, Switzerland
Christine Ronnewinkel, Germany
Elizabeth Rosenzweig, USA
Paul Sherman, USA
Ben Shneiderman, USA
Christian Sturm, Germany
Brian Sullivan, USA
Jaakko Villa, Finland
Michele Visciola, Italy
Susan Weinschenk, USA

HCI International 2013

The 15th International Conference on Human–Computer Interaction, HCI International 2013, will be held jointly with the affiliated conferences in the summer of 2013. It will cover a broad spectrum of themes related to human–computer interaction (HCI), including theoretical issues, methods, tools, processes and case studies in HCI design, as well as novel interaction techniques, interfaces and applications. The proceedings will be published by Springer. More information about the topics, as well as the venue and dates of the conference, will be announced through the HCI International Conference series website: http://www.hci-international.org/

General Chair
Professor Constantine Stephanidis
University of Crete and ICS-FORTH
Heraklion, Crete, Greece
Email: cs@ics.forth.gr

Table of Contents – Part I

Part I: DUXU Theory, Methods and Tools

Part II: DUXU Guidelines and Standards

Part III: Novel DUXU: Devices and Their User Interfaces

Part IV: DUXU in Industry

Part V: DUXU in the Mobile and Vehicle Context

Table of Contents – Part II

Part I: DUXU in Web Environment

Part II: DUXU and Ubiquitous Interaction / Appearance

Part III: DUXU in the Development and Usage Lifecycle

Part IV: DUXU Evaluation

Part V: DUXU beyond Usability: Culture, Branding, and Emotions

Conformity with User Expectations on the Web: Are There Cultural Differences for Design Principles?

Andreas Auinger[1], Anna Aistleithner[1], Harald Kindermann[1], and Andreas Holzinger[2]

[1] Upper Austria University of Applied Sciences, Department for e-Business, Steyr, Austria
[2] Medical University Graz, Research Unit HCI4MED, Institute of Medical Informatics, Statistics and Documentation, Austria
{andreas.auinger,anna.aistleithner, harald.kindermann}@fh-steyr.at, a.holzinger@hci4all.at

Abstract. User-centered Web design essentially impacts a website's success and therefore directly or indirectly influences a classic or digital enterprise's prosperity. "Conformity with user expectations" as one of seven dialogue principles according to the ISO 9241-110 standard is one critical success factor as it regards efficient and effective task completion. Over the past ten years, numerous recommendations for designing Web elements have been published, and some of them deal with conformity of user expectations. However, there are cultural differences concerning how design principles should be applied on Web elements. In this paper, we outline examples of their implementation, followed by discussing the results of an eye tracking study, which indicates that not all recommendations for design principles provided in related work – especially from the Anglo-American area – are valid for European end users and, finally, that their validity may change over time.

Keywords: cultural differences, conformity, user expectation, eyetracking, intercultural design principles.

1 Introduction

The effective, efficient and user satisfying implementation of the seven dialogue principles, i.e. suitability for the task, self-descriptiveness, conformity with user expectations, suitability for learning, controllability, error tolerance, suitability for individualization; represent the ISO 9241-110 standard's core concept for user centered web design.

User interfaces that conform with user expectations are defined as "*predictable user concerns, emerging from the applicational context and in accordance with generally accepted conventions*".

However, when people interact with websites, the standard does not include the extent to which conformity with user expectations contributes to task efficiency.

The following paper outlines the methodology and results of an eye tracking study that compares different web designs which were implemented according to specific design principles and corresponding design recommendations. Thus, design recommendations,

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 3–12, 2011.
© Springer-Verlag Berlin Heidelberg 2011

taken from pre-defined references, were tested on their qualification to allow a prompt orientation on different websites.

In corresponding usability literature, conformity with user expectations on websites is a common issue. Tirapat and Achalakul (2006) examined different hyperlink positions and were able to prove that users have very specific expectations on the position of web elements that need to be considered in web design. Cox and Fisher (2009) experienced, that confirmed user expectations, lead to an improved interaction quality that correlates positively with user satisfaction. Furthermore, Zhang and von Dran (2002) showed that users' expectations and their QoS criteria may vary according to different domains.

The following references provided the necessary design recommendations to prepare the study:

Nielsen and Tahir (2002), Jacobsen (2005), Nielsen and Loranger (2006), Shneiderman and Leavitt (2006), Cappel and Huang (2007), Angeli and Kundler (2008) and the European Committee for Standardisation (2006).

Subsequently, four exemplary web elements and their design recommendations have been chosen in order to conduct the eye tracking study and have been summarized in Table 1.

Table 1. Exemplary design recommendations

Web-Element	Design conformity with user expectations	Reference
I: Page Title	• Title should be unique and significant • Title should represent a site's content the best possible	Jacobsen, 2005
II: Logo	• To be positioned in the top left section of the page • Should be a hyperlink that leads back to the homepage	Shneiderman/Leavitt, 2006
III: Main Navigation **a) Link Title**	• Site content should be described as short and explicit as possible • No unnecessary additions • No repetition of the company's name • No addition of 'e-' or 'internet-'	Nielsen/Loranger, 2006
	• Target group should be familiar with the words used in a specific context	European Commitee for Standardization, 2006
b) Link Design	• Font color blue and underlined • Font color magenta after first visit	Cappel/Huang, 2007
	• Navigation menu has to be positioned on the left hand side	Shneiderman/Leavitt, 2006
c) Navigation tabs	• Designed like register menu of tabs • No simple bars with mouseover animation	Shneiderman/Leavitt, 2006
d) Navigation in Webshops	• Shopping cart-button (hyperlink) on the top right, below the basket's product list	Angeli/Kundler, 2008 Nielsen/Loranger, 2006
IV: Search Function	• Positioned in one of the top corners • White input field without any text • Describe the button with accurate terms such as 'search' or 'find' instead of 'go'	Nielsen/Tahir 2002 Nielsen/Loranger, 2006

Since eye tracking has been repeatedly approved as a reliable method in many studies (Duchowski, 2007), it is also appropriate in usability studies (Nielsen & Pernice, 2010). Furthermore, existing guidelines for usability need continuous reassessment with eye tracking technology (Cooke, 2004). Thus, eye tracking has been chosen for this study as the most reliable tool to measure efficiency.

2 Background and Related Work

User satisfaction is a very important aspect (Hassenzahl, 2001) (Stickel et al., 2009) and previous research indicates that the expectations of end users play a crucial role in their satisfaction. According to Tesch et al. (2005) this includes expectations regarding the skill levels exhibited by the providers of the IS services and products. Expectations are basically examined as gaps from perceived performance or as gaps from realistic expectations and although the interaction of these gaps has not been thoroughly researched, recent theories anticipate both gaps are crucial in meeting the desires of the end users. Previous studies indicate that expectations should be managed to higher levels and commonly understood across the end user and provider groups (Tesch et. al, 2005).

Interface consistency has been studied for quite a long time, actually since Graphical User Interfaces (GUIs) began to be used widely. Working with such user interfaces will be more efficient under the premise that a worker who is able to predict, what the system will do in any given situation, and can rely on the given rules (Nielsen, 2001). Consequently, the focus of research was on worker's productivity in order to achieve higher throughput and fewer errors. As a result of this goal, most early studies were on job performance of office workers, i.e. error rate and time to perform a task. The latter is the typical Human–Computer Interaction (HCI) approach and is usually considered in a transfer paradigm in which: *the higher the similarity between two tasks is, the higher the transfer, hence the consistency* (Tanaka, Eberts & Salvendy, 1991).

However, a strict establishment of the primary places of where consistency is most necessary is difficult. Grudin (1989) separated consistency into *internal* interface consistency and *external* interface consistency, wherein internal refers to consistency within a task and external means consistency among various tasks. Ozok & Salvendy (2000) classified it into three sub types, establishing the *three-dimensional model of interface consistency:*

(i) Conceptual consistency (language, stereotypes, task concept, skill transfer, output consistency, hierarchical order of concept, etc.);
(ii) Communicational consistency (moving between screens, menus, user conventions, between-task consistency, distinction of tasks and objects, etc.);
(iii) Physical consistency (color, size, shape, location, spacing, symbols, etc.).

Ad i) *Conceptual consistency* can be defined as the consistency of metaphor applied to an interface feature or an action that is embodied within a feature. Frequent and inconsistent use of synonyms, instead of using the same words for the same items, is unhelpful. Leaving something to students' conception and interpretation due to the lack of explicitness is also regarded as conceptual inconsistency (Grudin, 1989), (Ozok & Salvendy, 2000).

Ad ii) *Communicational consistency* can be defined as the consistency of both input and output of the interface. It deals with how the user interacts with the computer interface and whether the means of interaction are consistent for fulfilling the same or similar tasks.

Ad iii) *Physical consistency* can be defined as the consistency of the visual appearance of an interface feature and indicates that the features are supposed to be consistent with the users' mental models (Satzinger, 1998).

Although this has been known for quite a long time, research on the relationship between *consistency and human work and learning processes* has only recently been documented, and Satzinger & Olfman (1998) pointed out that very few studies have investigated the effects of interface consistency on work and learning performance. To design an appropriate user interface demands insight into the *behaviour of the end users* and the application of user centered development (Norman & Draper, 1986), (Holzinger, 2002), (Norman, 1986), in order to achieve a true interaction. This is of essential importance, since working with interactive media is generally highly demanding from the perspective of the limited cognitive processing capabilities of the end users (Holzinger, Kickmeier-Rust & Albert, 2008), (Holzinger et al., 2009). Daily practice shows that many end users have difficulty working with electronic systems, since they are often unable to form a mental model of the system and their current position within its complexity. However, when striving for a design following the "principle of the least surprise", we are faced with the problem that designers and developers are rarely able to predict exactly what the end users really expect (remember Steve Krug (Krug, 2000): "Don't make me think!").

In addition, related work in "*cultural usability*" (cf. Vatrapu & Suthers 2010) strongly focuses on the relevant aspects of user interface design:

(i) *Cultural conventions*: Research in the 1990's focused on localization and internationalization of user interfaces with respect to languages, colors and convention of data, time and currency (Fernandes, 1995, Khaslavsky, 1998; Russo & Boor, 1993).

(ii) *Cultural influences* on usability evaluation methods and usability processes were found in the usability assessment methods of focus groups (Beu, Honold, & Yuan, 2000), think-aloud (Clemmensen, Hertzum, Hornbæk, Shi, & Yammiyavar, 2009; Yeo, 2001), questionnaires (Day & Evers, 1999), and structured interviews (Vatrapu & Pérez-Quiñones, 2006).

(iii) *Cultural differences* with respect to usability processes were found in the understanding of metaphors and interface design (Day & Evers, 1999; Evers, 1998).

(iv) *Cultural Web design:* Finally, culture was found to affect web design (Marcus & Gould, 2000), objective and subjective measures of usability (Herman, 1996), and subjective perceptions and preferences in mobile devices (Wallace & Yu, 2009).

3 Methodological Design

The study was conducted with 22 test subjects from Austria between the ages of 17 and 28. The probands use the internet daily to several times a week, for business as well as private matters. They were divided into two research groups of 10 and 12 subjects. Hence, each research group was shown one of two different layout designs containing specifically designed web elements; one of each was designed to conform to the chosen design recommendation concerning the design principle "conformity with user expectations"; one was not (test 1 to 4).

The goal was to complete a task by using those pre-defined design elements/ hyperlinks positioned on the web site and to confirm them by a mouse click to finish the particular task (see table 2).

Apart from the different positions and designs of the Web elements, respective design principles to be tested, the websites were completely identical. The eyetracking data was recorded by a 120Hz eyegaze eye tracking system from "Interactive Minds". It enabled a comparison of the proposed Web site alternatives regarding general indicators and ratios, as well as required times for task completion and hit rates.

Table 2. Tested Web elements

	Web-Task	**Alternative 1**	**Alternative 2**
Test I: **Position of the <u>Main Navigation</u>**	Find the hyperlink "Product Overview" as a part of the main navigation	Positioned on the left hand side → ***conforming*** to user expectations	Positioned on the right hand side → ***not conforming*** to user expectations
Test II: **Position of the <u>Search Function</u>**	Find the search function and click on the search-button	Top left → ***conforming*** to user expectations	Middle right → ***not conforming*** to user expectations
Test III: **Position of the <u>Website Logo</u>**	Find a link back to the homepage without clicking the "Homepage"-button or the browser's "Back"-button	Top left → ***conforming*** to user expectations	Top right → ***not conforming*** to user expectations
Test IV: **Design of <u>Navigation Tabs</u>**	Find the hyperlink "About TopZoo" as a part of the navigation tabs	Tabs look like real-world tabs → ***conforming*** to user expectations	Bar with hyperlinks and mouseover animation → ***not conforming*** to user expectations

4 Results

In order to be able to calculate the relevant indicators, areas of interest were defined with the eyetracking analysis software Nyan 2.3.5 for each web site. Additionally, it was possible to calculate the average time to finish each task.

Test 1: Position of the Main Navigation. In comparison, the time to the first fixation of the main navigation took about 0.84s longer with the alternative 2 (alt1:1.56s | alt2:2.4s). The average fixation duration for the main navigation (alt1:0.49s | alt2:0.6s) and the product overview (alt1:0.47 | alt2:0.66s) was nearly the same for both alternatives. Both times, alternative 1 was fixated more often and more securely (alt1:90% | alt2:81.8%). Furthermore, the test subjects using alternative 1 could finish the task 0.89s faster (alt1: 4.73s | alt2: 5.59s). Thus, alternative 1, which was designed according to user expectations, allows a more efficient handling of the task.

Test 2: Position of the Search Function. Results show that the time to first fixation of the search function took place 2.3s *sooner in alternative* 1 (alt1:1.35s | alt2:3.6s). The average fixation duration differs with 0.75s (alt1:2.25s | alt2:3s). Greater differences appeared in the task completion time, where alternative 1 with its search function on the bottom left was completed 2.79s more quickly (alt1:4.22s | alt2:7.01s). Figure 1 shows exemplary heat map screenshots of the implemented design prototypes for testing the search function.

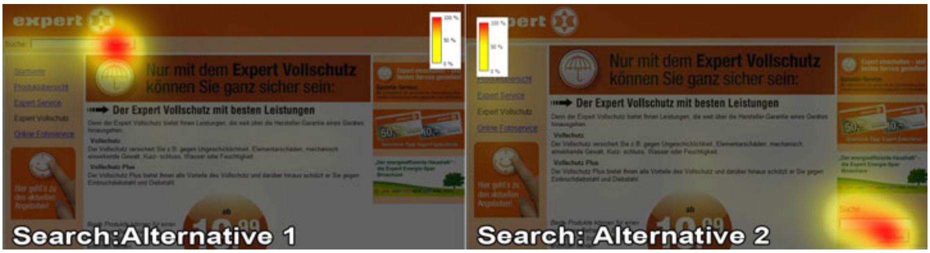


Fig. 1. Heatmaps of test 2 “Position of the search function”

Test 3: Position of the Website Logo. Only three people in each group used the website-logo as a backlink function (Shneiderman/Leavitt, 2006) within an acceptable period of time. The average time to the first fixation in *alternative 2* happened more quickly than in alternative 1 (alt1:5.31s | alt2:4.26s), however the fixation duration (alt1:1.62s | alt2:0.83s) as well as the number of fixations (alt1:6 | alt2:2) proved to be more reliable in alternative 1. In view of the large number of people (75%) who were not able to finish this task and the longer thinking time, it can be concluded that the *backlink function* of website logos, according to the literature (Shneiderman & Leavitt, 2006), *is not very well-known* within the tested probands group from Austria.

Test 4: Design of Navigation Tabs. As outlined in table 2, alternative 1 shows register menus of tabs while alternative 2 makes use of a simple bar with mouseover animated hyperlinks. A comparison shows that the fixation duration (alt1:1.29s | alt2:1.07s) and the number of fixations (alt1:3 | alt2:2.7) is better in alternative 2. Also the time to the first fixation (alt1:4.14s | alt2:2.22s) as well as the time to finish the overall task (alt1:14.82s | alt2:10.53s) shows better results in alternative 2. In conclusion, the implementation of navigation tabs as proposed by Shneiderman and Leavitt (2006) does not comply with the tested probands expectations.

Are there Cultural Differences? Effective image design on websites is central for successful international communication (Amant, 2005). The results of tests 3 and 4 leave room for interpretation. The implemented design recommendations of tests 3 and 4 were proposed by Shneiderman and Leavitt (2006). They did not work persuasively with the conducted 22 test subjects from Austria. Are there cultural differences in how to implement specific design principles and how the probands use them in Austria, which should be valid for the German-speaking Europeans, in contrast to the Anglo-American room?

Table 3. Results of the Eyetracking Study

	Test 1				Test 2	
	AOI: Main Navigation		AOI: Link "Product Overview"		AOI: Search Function	
	alt.1	*alt. 2*	*alt. 1*	*alt. 2*	*alt. 1*	*alt. 2*
Time to First Fixation [s]	1.56	2.4	1.92	2.71	1.35	3.64
Fixation Duration [s]	2.44	2.07	0.95	1.57	2.25	3
Ø Fixation Duration [s]	0.49	0.6	0.47	0.66	0.48	0.5
AOI Hit Rate [%]	90.0	81.8	90.0	81.8	90.0	90.9
Number of Fixations	5	3.5	2	2.4	4.7	6
Task Completion [s]			4.73	5.59	4.22	7.01

	Test 3		Test 4			
	AOI: Website-Logo		AOI: Tab Navigation		AOI: Link "About TOP/ZOO"	
	alt. 1	*alt. 2*	*alt. 1*	*alt. 2*	*alt. 1*	*alt. 2*
Time to First Fixation [s]	5.31	4.26	2.09	1.02	4.14	2.22
Fixation Duration [s]	1.62	0.83	4.52	2.68	1.29	1.07
Ø Fixation Duration [s]	0.27	0.41	0.35	0.3	0.43	0.4
AOI Hit Rate [%]	100.0	66.7	100.0	90.0	91.0	70.0
Number of Fixations	6	2	13	8.9	3	2.7
Task Completion [s]	8.77	9.86			14.82	10.53

Ad test 3: In test 3, only 6 of 22 probands used the backlink function in an appropriate period of time. Accordingly, 16 probands did not use or even know about this functionality.

Ad test 4: In the case of test 4, it apparently seems, that German-speaking Europeans are not used to explicit register tabs for menu design in the same dimension that people in the Anglo-American room are. Shneiderman and Leavitt (2006) recommended the following menu tab design example in their "Research Based Web-Design & Usability-Guidelines" (Fig. 2):

Fig. 2. Recommended register menus (Schneiderman & Leavitt, 2006)

The probands located the wanted menu option during both test alternatives in an appropriate period of time but were significantly faster in completing the task with alternative 2 (alt1:14.82s | alt2:10.53s). They seemed to feel uncertain and were not used to register menus and obviously have adopted modern menu design elements with dynamic menu options and mouseover effects in the German-speaking room, which were explicitly not recommended as design principles.

Hence, design principles are not generalisable (as proposed by Shneiderman and Leavitt 2006) and seem to depend on the intercultural context. To overcome this deficiency, design principles and the corresponding recommendations, need to be evaluated in different cultural environments all over the world and validated explicitly concerning the cultural context where they are intended to be implemented.

5 Discussion and Conclusion

The four tested design recommendations for web elements to fulfil the design principle "conformity with user expectations" show considerable differences concerning adequateness of task completion.

The results for the typical position of the main navigation and the search function show compliant operating speed and confirm the recommended design suggestions for conformity with user expectations. The website logo as a back-link function has surprisingly been used only by a very small number of sophisticated internet users.

The suggestion to design register menus of tabs seems to be out-of-date and dependent on the cultural area where it is applied, as the indicators showed a completely contrary result. We may interpret, that German-speaking users have rather adopted their expectations to dynamic menu designs with mouseover effects, and they are able to use them in an efficient and effective way.

In conclusion, there will always be cultural differences, necessary adaptations of design and colour and design adaptations according to a specific target group that have major priority for accurate implementation. The validity of common recommendations needs to be continuously reassessed and evaluated for their specific cultural context, as the dialog principle "conformity with user expectations" may change in the course of time and may be interpreted dependent on cultural differences.

References

1. Aman, K.S.: A Prototype Theory Approach to International Image Design. IEEE Transactions on Professional Communication 48(2) (2005)
2. Angeli, S., Kundler, W.: Der Online-Shop/Handbuch für Existenzgründer, München (2008)
3. Beu, A., Honold, P., Yuan, X.: How to Build Up an Infrastructure for Intercultural Usability Engineering. The International Journal of Human-Computer Interaction 12(3&4), 347–358 (2000)
4. Cappel, J.J., Huang, Z.: A Usability Analysis of Company Websites. Journal of Computer Information Systems, 117–123 (Fall 2007)
5. Clemmensen, T., Hertzum, M., Hornbæk, K., Shi, Q., Yammiyavar, P.: Cultural cognition in usability evaluation. Interacting with Computers 21(3), 212–220 (2009)
6. Cooke, L.: Improving Usability Through Eye Tracking Research. In: Proceedings of Professional Communication Conference (2004)
7. Cox, A., Fisher, M.: An Expectation-Based Model of Web Search Behaviour. In: Proceedings of the Second International Conferences on Advances in Computer-Human Interactions (2009)
8. Day, D., Evers, V.: Questionnaire Development for Multicultural Data Collection. Paper presented at the International Workshop on Internationalisation of Products and Systems, IWIPS-1999 (1999)
9. Duchowski, A.T.: Theory and Practice Eye Tracking Methodology, 2nd edn. (2007)

10. European Commitee for Standardisation. DIN EN ISO 9241-110:2006/Ergonomie der Mensch-System-Interaktion/Teil 110: Grundsätze der Dialoggestaltung (2006)
11. Evers, V.: Cross-cultural Understanding of Metaphors in Interface Design. Paper presented at the Attitudes toward Technology and Communication, London (1998)
12. Fernandes, T.: Global Interface Design. Academic Press, London (1995)
13. Grudin, J.: The Case against User Interface Consistency. Communications of the ACM 32(10), 1164–1173 (1989)
14. Hassenzahl, M.: The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction 13(4), 481–499 (2001)
15. Holzinger, A.: User-Centered Interface Design for Disabled and Elderly People: First Experiences with Designing a Patient Communication System (PACOSY). In: Miesenberger, K., Klaus, J., Zagler, W.L. (eds.) ICCHP 2002. LNCS, vol. 2398, pp. 33–41. Springer, Heidelberg (2002)
16. Holzinger, A., Kickmeier-Rust, M., Albert, D.: Dynamic Media in Computer Science Education. Content Complexity and Learning Performance: Is Less More? Educational Technology & Society 11(1), 279–290 (2008)
17. Holzinger, A., Kickmeier-Rust, M.D., Wassertheurer, S., Hessinger, M.: Learning performance with interactive simulations in medical education: Lessons learned from results of learning complex physiological models with the HAEMOdynamics SIMulator. Computers & Education 52(2), 292–301 (2009)
18. Khaslavsky, J.: Integrating Culture into Interface Design. Paper presented at the Conference on Human factors in computing systems (CHI 1998), Los Angeles, USA (1998)
19. Krug, S.: Don't Make Me Think: A Common Sense Approach to Web Usability, Indianapolis (IN), New Riders (2000)
20. Marcus, A., Gould, E.M.: Cultural Dimensions and Global Web User-Interface Design. Interactions 7(4), 32–46 (2000)
21. Nielsen, J.: Coordinating User Interfaces for Consistency. The Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann, San Francisco (2001)
22. Nielsen, J., Loranger, H.: Web-Usability. Addison-Wesley, Munich (2006)
23. Nielsen, J., Pernice, K.: Eyetracking Web Usability. Pearson, Berkeley (2010)
24. Nielsen, J., Tahir, M.: Homepage Usability/50 Websites Deconstructed. New Riders, Munich (2002)
25. Norman, D.A.: Cognitive engineering. In: Norman, D., Draper, S. (eds.) User Centered System Design: New Perspectives on Human-Computer interaction. Lawrence Erlbaum, Mahwah (1986)
26. Norman, D.A., Draper, S.: User Centered System Design. Lawrence Erlbaum, Mahwah (1986)
27. Russo, P., Boor, S.: How Fluent is your Interface? Designing for International Users. In: Paper presented at the Conference on Human factors in computing systems (CHI 1993), Amsterdam, The Netherlands (1993)
28. Satzinger, J.W.: The effects of conceptual consistency on the end user's mental models of multiple applications. Journal of End User Computing 10(3), 3–14 (1998)
29. Satzinger, J.W., Olfman, L.: User interface consistency across end-user applications: the effects on mental models. Journal of Management Information Systems 14(4), 167–193 (1998)
30. Shneiderman, B., Leavitt, M.O.: Research Based Web-Design & Usability-Guidelines (2006), `http://www.usability.gov/guidelines/` (April 5, 2010)

31. Stickel, C., Ebner, M., Steinbach-Nordmann, S., Searle, G., Holzinger, A.: Emotion Detection: Application of the Valence Arousal Space for Rapid Biological Usability Testing to Enhance Universal Access. In: Stephanidis, C. (ed.) UAHCI 2009. LNCS, vol. 5614, pp. 615–624. Springer, Heidelberg (2009)
32. Tanaka, T., Eberts, R.E., Salvendy, G.: Consistency of Human-Computer Interface Design - Quantification and Validation. Human Factors 33(6), 653–676 (1991)
33. Tirapat, T., Alchalakul, T.: Usability Assessment for Hyperlink Methods. In: Proceedings of the International Conference on Hybrid Information Technology (2006)
34. Ozok, A.A., Salvendy, G.: Measuring consistency of web page design and its effects on performance and satisfaction. Ergonomics 43(4), 443–460 (2000)
35. Vatrapu, R., Pérez-Quiñones, M.: Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. Journal of Usability Studies 1(4), 156–170 (2006)
36. Vatrapu, R., Suthers, D.: Intra- and Inter-Cultural Usability in Computer Supported Collaboration. Journal of Usability Studies 5(4), 172–197 (2010)
37. Wallace, S., Yu, H.: The Effect of Culture on Usability: Comparing the Perceptions and Performance of Taiwanese and North American MP3 Player Users. Journal of usability Studies 4(3), 136–146 (2009)
38. Yeo, W.A.: Global-Software Development Life Cycle: An Exploratory Study. In: Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems, pp. 4–11 (2001)
39. Zhang, P., von Dran, G.M.: User Expectations and Rankings of Quality Factors in different Web Site Domains. International Journal of Electronic Commerce 6(2), 9–33 (2002)

A Philosophical Approach about User Experience Methodology

Marcos N. Beccari and Tiago L. Oliveira

Departamento de Design , Universidade Federal do Paraná, Rua João Ramalho, 302 - casa07, Curitiba, 82820280, Brazil
contato@marcosbeccari.com, etsete@gmail.com

Abstract. The purpose of this paper is to identify some of the possible contributions of the entitled Philosophy of Design to the processes involved in the User Experience methods. After a brief introduction on User Experience principles and methods, we will make a brief overview of the history of research in Design. Moving on we shall review some of the main precepts of Philosophy of Design and, finally, make evident the scientistic and pragmatic predominance of the User Experience methods.

Keywords: User Experience Methodology, Philosophy of Design, Post-Positivism, Terence Love.

1 User Experience Methods

According to ISO 20101, User Experience involves human perceptions and responses that result from the use (or the anticipation of use) of a product, system or service. The User Experience includes all emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and achievements that occur before, during and after the use of a product. Bevan (2009) shows that the difference between User Experience and Usability is not exactly the measurements or methods used, but on the emphasis and focus established during the development of a product. User Experience points to a more global projective goal: not just attain effectiveness, efficiency and satisfaction, but it aims to enhance the entire experience of the user, from the expectation, through interaction and finally the reflection about the experience (op. cit.).

In a practical sense, User Experience probes to get closer to the user as a human being and not just the limited situation such as facing a particular product. Instead of concerning only in identifying and correcting problems, methods in User Experience question about what people do and why they do it. Still, there are methods in common with Usability, among these: Longitudinal Studies, Ethnography, Contextual Inquiry, Interview Retrospective, Focus Group and Daily. However the goal of this work is restricted only on some specific methods or techniques of User Experience, collected and described by Law et. al. (2009).

The first to be mentioned are called Emocards. It consists in 16 cards containing facial expressions, distributed in points along a circle divided in four parts: intense unpleasant, intense pleasant, calm unpleasant and calm pleasant. Crossing the circle, there are scales of values of satisfaction and provocation which are assigned to the

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 13–22, 2011.
© Springer-Verlag Berlin Heidelberg 2011

product. After a simple survey, its calculated the number of times each value was mentioned. Similarly, the method "PrEmo" uses animation to represent positive and negative emotions. The graphic animation of the body language combined with facial expression provides basic details on the representation of an emotion. After collecting data, results are compares and a final graph is generated.

As for the photographic language, the Moodboards evaluate, by means of questionnaires, the mood, perception and one's desires during the interaction with a product. For example, three different pictures form alternatives to answer the question "Which images best represents the environment where you want to use this product?".

Presenting with more options, Reaction Cards are a set of 118 adjectives or phrases that are selected by users to describe their interaction with the product and their perception of its qualities. When it comes to digital monitoring of the users interaction with a system, the instrument "True" is largely used: one records all interaction and intercalates it with contextual questions in certain periods of time. The behavioral data are then compared with data regarding users attitudes towards the system. Similarly, though focussing on the evaluation of online systems, Attrakdiff is an instrument based on a semantic differential with a 7 point scale: the user selects a position that can go from "confusing" to "clearly structured", for example.

In the face of these and many other examples of User Experience methods, how has the designer been acting? Which have been the foundations of their methodology? Does Design has its own body of knowledge? These are the kind of reflections that permeate this work. It is easy to see that User Experiences methods and techniques as a project discipline, even having its singularities, does not make clear what are the epistemological foundations on which they stand, making clear only the objectives they seek. There is no concern with the knowledge that precedes such methods - being only questionnaires that long to turn subjective facts into something objective strictly in order to "make it work". Our attempt here is to emphasize the pragmatic nature of the User Experience methods, so we will no longer simply define concepts as we have done so far, our attempt is to now open paths that starts from a philosophical attitude and points to a plural understanding of User Experience as a theoretical field. With this in mind, we will not break away completely from the methodological modulation of User Experience, but on the other hand, we wish to demonstrate how this modulation becomes fragile and perplexed when the perspective for the analysis is Philosophy of Design..

2 Historical Overview of Design Research

Before we analyze the existing epistemological paradigms in the User Experience approach, we shall take this opportunity to briefly raise some of the historical premises regarding Design Research that are essential to understand these paradigms. A context in which the modern discourse prevailed in Europe, form follows function or functionalism was the dominant tenet for decades in architecture and Design (FONTOURA, 1997). According to Cross (2007), Design Research initiated only with the first Conference on Design Methods, held in London in 1962. In an attempt to consolidate Design methodology as a scientific discipline, the Design Methods movement sought to replace the intuitive process, still recurrent at the time, for the application of purely scientific and rational procedures, operational research methods, management techniques of decision making, creativity techniques, etc.

If in the 1960s the motto was to standardize and rationalize, the 70s was marked by the rejection of scientific Design methods of (op. cit.). However, the second generation of researchers, drawing on the pragmatist ethics of reasonableness, facilitated applied research and maintained the scientific status of Design. This made possible, in the 80s, a period of solid achievement of Design Research with the emergence of Design Research scientific journals2. Opposed to the pragmatism of the Design Engineering movement, the Italian Il Nuovo Design3 wanted the return of the inductive method by means of a a paradigm shift process in the sense give by Thomas Kuhn, in the project methodology. It was a critical and reformist position that followed a distinct path from the emerging Design Think, a third strand that wanted to study Design on its own, that is, based on the premiss that Design has its own objects of study and its own ways of knowing them.

It is only that in the 90s, as Cross (2007) points out, that Design would reach its adulthood with a significant expansion of conferences and journals. On the other hand, it is in this context that emerges the recurrent Philosophy of Design, a movement that criticizes the fragmented tendency of confusion, fusion and multiplicity in Design Theory.

3 Philosophy of Design

"That there exists a substantial amount of confusion with respect to the underlying basis of many theories, concepts and methods. That in developing and validating theoretical aspects of the study of design, many writers are unjustifiably conflating concepts drawn from a range of sources. That there exists an unnecessary multiplicity of design theories and concepts. That the terminology of design research has become unnecessarily and unhelpfully confused and imprecise by dint of the above points" (LOVE, 2000, p. 295)

Love had noticed this since 1992, in attempts to collect the main theoretical terms in the Design research literature in his publication Social, Environmental and Ethical Factors in Engineering Design Theory: a Post Positivist Approach (LOVE, 1998). After observing that there are so many different variations of Design and Design Processes as there are authors who write about them, different authors that use the same words for different concepts or refer to the same concepts with different words, Love (1998) acknowledged in his cataloging proposal an almost impossible task. Eder (1981) had already listed several words that were used with different meanings in Design Research, concluding that their meaning depended solely on the cultural context in question. Thus, the problem of internal confusion in Design Research became recurrent in the early 90s, and facing an increasing conceptual indeterminacy of Design, many researchers have proposed countless solutions or ways to deal with it. Some sought only to critically examine the various existing concepts, others preferred to organize and systematize what they considered to be Theory of Design.

Adopting Love's (2000) definition of Philosophy of Design, we will list here some authors that also refer to a philosophy of Design to understand its true meaning and shape that User Experience takes as a result of that. Remembering that Design Theory is the discipline that investigates Design Methodology, Design Methods, Design History, among other studies that attribute a role, a consistency and validity to Design as an area of research (op. cit.). Philosophy of Design, on the other hand, would be a discipline apart that tries to investigate Design Theory as its object of study, and can also be considered a meta-theory (op. cit.).

This new discipline was initially proposed by some collaborators of the international journal Design Studies. Terence Love (2000) elaborated an analytical and meta-theoretical approach for the construction of a unified theoretical body to Design, and two years later, published the article Constructing a coherent cross-disciplinary body of theory about designing and designs: some philosophical issues4 (LOVE, 2002) in which some of the key elements are proposed for a more coherent grounding to research and construction of theories in Design. Noticing a growing number of researches5 interested in developing a unified theoretical body of Design, Love describes the inconclusive increase of theoretical conflicts between researchers in different contexts. To Calvera (2006, p. 100), "the situation can be seen as the fight of a young discipline against his parents, older and sometimes conservative. "Such conflicts will increase the difficulty of theoretical validating that takes account epistemological and ontological aspects - which could define according to Love (2002), a consistent body of Design Theory. The lack of clarity on the focus, range and limits of existing theories generates more obstacles every year for initiating researchers (graduate students) who seek to establish a satisfactory literature review for their research, not being able to identify the epistemological foundations facing a wide variety of theoretical perspectives (op. cit.).

"*The analyses in the paper point to a significant political question to be resolved by the field: whether researchers who have an investment in past literature with its philosophically problematic foundations and domain-specific theories can be persuaded to support the development of new and more coherent cross-disciplinary foundations and the building of a single body of theory and knowledge about designing and design*s." (LOVE, 2002, p. 359).

Before investigating exactly how Design researchers are merging many concepts inappropriately, Love (2000) explains how these mergers may occur and how they can lead to the incorrect development of Design theories: If A is related to B and B is related with C, then A is related to C. Initially it is a mistake in the relationship between abstract entities - "The cat has four legs and a dog has four legs therefore a cat is a dog" (LOVE, 2000, p. 297). Although it recurs to the primary socratic rhetoric6, the intention here is to show how logical reasoning can lead us to inappropriate associations that, once taken as theoretical grounds, can generate naively equivocated conclusions, or fallacies. For instance: "Designers think and cognitive psychologists study thinking, therefore, research into design lies within the discipline of Cognitive Psychology" (op. cit.). As much as this type of deduction are not made in a such naive way in Design, Love's distrust points to the uncontrolled production of literature related to Design research since the 50s. Even with hundreds of books and articles published every year, theories are speculative proposals coming from a wide variety of theoretical perspectives, however, they are not sufficient to make an epistemological and ontological analysis (op. cit., p. 298). This is because there is little agreement on fundamentals aspects (such as the definition of the word Design), but this does not prevent research to continue to develop in a fragmented way in various theoretical provinces (op. cit.).

In this context, the misguided and naive associations become inevitable in a theoretical development with a temporal base, that is, as attempts are made to describe patterns in theories made by previous generations (op. cit.). One of the first authors to propose a "simplification in the research paradigm of Design" was Cross (1984) when trying to map some of the themes that were being developed sequentially, each assigned to correct the failures of its predecessor. In other words, Cross (op. cit.) identified some chronological paradigms that use the same terms and concepts to refer to different meanings according to each historical context. The thematic thread of these paradigms goes to the following direction: management of the Design process, the structure of Design problems, the nature of the Design activities; reflections on the fundamental concepts of Design (op. cit.). Love (2000) suggests the addition of two other themes to this developmental paradigm: knowledge about the environment in which Design occurs, and the necessary knowledge to conceive Design, that is, about objects and Design processes7. Ten years after his paradigmatic analysis, Cross (1993) abandons his kuhnian approach and recognizes such naive mergers described by Love in a "conflactionary" development of theories that crosses through all paradigms simultaneously, thus creating a terminology confusion. Dixon (1988), on the other hand, confronts this terminological confusion by arguing that research in Design still is in a pre-theoretical stage, that is, only with the purpose of establishing "testable" scientific theories to Design, and all attempts established so far, however, are impossible to be applied in practice (op. cit.).

Anyway, it is not hard to notice a post-positivist trend among most researchers, including Love, that proposes a philosophical approach to the theories of Design. Mainly represented by Karl Popper and Thomas Kuhn, post-positivism (also called post-empiricism) is a philosophical movement that seeks to criticize and improve positivism in a meta-theoretical way (ZAMMITO, 2004). Following Poppers principle of falsifiability, design philosophers declare that is impossible to verify whether a theory is true, although it is possible to identify false theories when they are arranged in a manner favorable to refutation. Kuhn's perspective is adopted with the conception that not only individual theories, but all world-views should change in response to evidence presented in each cultural-historical paradigm. This argument is defended, for example, in Margolins (1992) on investigating the foundations of the cultural basis behind Design History. Recurrent in the social sciences for practical and conceptual reasons, post-positivism is not a pure form of relativism, because it still preserves many foundations from positivism: the ontological realism, the desire for objective truth and the use of experimental methodology (PHILIPS, BURBULES, 2000).

By following this path, Design philosophers believe that, generally, human knowledge is not based on incontrovertible truths, but only on hypotheses. If we build a brief timeline on the influence of post-positivist in the Philosophy of Design, we could begin the research with Thomas and Carroll (1979) who hypothesized that Design is best conceived with theories that value the individual and psychological dimension. As a deployment to this, Daley (1982) starts to examine the role of objects in a psychological dimension, while Dilnot (1982) notes that the limitations of the definitions of Design generally exclude the social context in which they operate. This motivated Broadbent (1984) to critically discuss the theories constructed in Design, directly influencing the investigations of Coyne (1990, 1991) that assumes an increasingly post-positivist character. A short time later, Sargent (1994) proposes for the first time a meta-theoretical discussion to argue about the impossibility of an unification between science and design, claiming the existence of an incommensurability of views on research in Design. In the same direction, Liddament (1999) is concerned with "coded" and "computational" nature of the research in Design as a limitation to the ontological, epistemological and methodology development of Design. Simultaneously, Galle (1999) explores how the definition of Design is dependent on the explanation of human agents in different interactive situations. Finally, it is also worth mentioning the new approach proposed by Oxman (1999) for a Design education focused on the dialectical nature of the word Design, associating cognitive theories with the post-positivist epistemology.

However, it is especially with Bamford (2002), in his article From analysis/ synthesis to conjecture/analysis: a review of Karl Popper's influence on design methodology in architecture, that post-positivist influence in philosophical research in Design peaked. "All designing (...) can reasonably be described as hypotheses or, in some respects, conjectures, given the usual unknowns and uncertainties." (BAMFORD, 2002, p. 260). While Bamford (op. cit.) recognize that Design is not literally made out of trial and error, his critical thinking goes against the analysis / synthesis that would be prevalent in the Design methods, that is, deductive reasoning. Instead, Bamford (op. cit.) proposes an inductive reasoning, especially by means of analogies. Returning to Schön's (1988) assumption that both guesswork and technical knowledge, ideas and algorithms are part of the problem solving process in Design, Bamford (2002) is positioned in a more unilateral way demonstrating reasons to reject the analysis / synthesis in favor of conjecture / analysis.

Bamford's (2002) article was published in the 23rd Volume of the journal Design Studies, being this issue specifically dedicated to the emerging issue of Philosophy of Design. In the article that closes this issue, Love (2002) complains that the word Design and its derivatives such as projects and the verb project are being used in a sense quite diluted in the literature. After postulating that all the key concepts in Design should be chosen more carefully, Love (op. cit.) proposes a particular conception of Design as being a primary function of man similar to thinking and feeling. In a similar vein, Houk, Vermaas, Dorst and Vries (2002) suggest a definition of Design as a specific type of action, something that involves plans, intentions and practical reasoning. With respect to the issues raised by Bamford (2002) on methodology of Design versus scientific methodology, Kroes (2002) considers the Design methodology extremely prescriptive and process-oriented, as opposed to classical scientific methodology, that he believes to be descriptively and product-oriented. Following this assumption, Kroes (op. cit.) believes we should establish some criteria for measuring quality, success and failure of the process in Design. Trott (2002) also accuses the lack of criteria quality in Design ("standards of excellence", as she calls it), suggesting the platonic maximum (types or properties of abstract thought and timeless, as entities existing independently of their instances) as a source to such criteria. A third philosophical approach to quality in Design is suggested by Baljon

(2002) that, through a systematic analysis of some historical evidence, seeks to elucidate the mechanisms of success of the "canons of Design". Interesting to point out that the approach of Baljon (op. cit.) is purely historicist, that is, it is considered that the history of Design is more important than any other theory.

In opposite to the normative criteria of quality for Design, as well as for the recurrent Popperian empiricism, Coyne, Park and Wiszniewska (2002) prefer to adopt a terminology of the phenomenological analysis, specifically Heidegger's notion of "revelation"8 applied to the Design process. Such an approach naturally gives emphasis to the artifacts produced in Design: on one hand they are physical objects with a certain structure on the other are also intentional objects with a given function. This "dual nature" of objects is also exploited by Kroes (2002) that, through an analogy with the principle of Niels Bohr9, posits that it is not clear, philosophically speaking, the relationship between function and structure of an artifact, particularly referring to the transition between one concept and another. Working with Kroes in a project in common, Houk, Vermaas, Dorst and de Vries (2002) mapped in detail a sequence of actions to specify both the physical structure of an object and the fulfillment of its function. Extending this issue to the collaborative Design framework, Bucciarelli (2002) contends that linguistic is what allows designers to fill the gap between function and structure. However, Bucciarelli (op. cit.) rejects the possibility of a strictly rational and instrumental method for reconciling the linguistic differences, since designers are satisfied with very limited linguistic. Similarly, Bestelius and Doevendans (2002) rejects any attempt to reconcile the multiplicity of views, although proposing an alternative to explore rhetoric as a way of undermining the prevailing pragmatism.

This predominance was identified after an historical analysis on the transition of "modernity" to "postmodernity" in the Design field, analysis made by this Bestelius and Doevendans (op. cit.) and Bamford (2002). Proposing paradigm shifts in patterns of thought in Design, both researchers report that the methods of Design are as antiquated as the scientific methods adopted in the seventeenth century (op. cit.). With this, we return to the meta-theoretical analysis, prevalent in almost all authors, including Baljon (2002) - in this case, the author adopts a historical meta-analysis as a philosophical explanation to history itself. This trend (if it can so be called) is against Love’s (2002) proposal for a unified theoretical body of knowledge about Design. It is "a foundation for research and theory in Design and a coherent and interdisciplinary theoretical body of knowledge that is not confused with other disciplines" (op. cit. p. 345). To this end, Love suggests a system of research areas to be further explored in Design, discusses the delimitation of some fields of knowledge close to Design and offers guidance for the definition of key concepts for building theories of Design. However, this proposal contradicts Bucciarelli’s (2002) empirical findings about the impossibility of a unified language for Design, just as it does not corroborates the apparent complementarity of physical concepts and intentional artifacts dictated by Kroes (2002). It can be noticed, therefore, a first point of disagreement between these authors, while some do not believe in the possibility of building a unified body of theory in Design, defending that the multiplicity of theoretical perspectives somewhat positive, others are concerned about how to develop such pretension.

4 Philosophical Questions on User Experience

In face of this brief overview, the following pertinent question is: has the whole taxonomy contained in our methods been working in User Experience? Can any result of the interaction process between user and product be considered consensual? And can all Design projects be viewed as a structured consensus? These are statements on which, despite being in full accord with the pragmatist logic, does not manifest a significant impact on the theoretical development of the discipline of User Experience. The issue can not be summarized only in "the user has the mere status of an isolated stage in the Design process, or should be simultaneously focused along with all stages?" because the problems brought by Philosophy of Design is in the process itself that is faced with the dilemma: "user experience as means or end?".

In the user experience as means to achieve a particular end, the end is constantly simulated in a mechanism of a "funnel", that is, it is not the ideas that already exist that must be controlled, but those that should not exist, excluding all those considered unacceptable. It is therefore a dialectical and manichean reasoning, in which occurs the overlap of a thesis and an antithesis in order to reach a synthesis previously stipulated. On the other hand, when the User Experience is seen as the goal, it presented, from the beginning, its end, as if it had always been like that, that is, an edifying intent based on an idea already built. In this case, everything we do due of something else is considered partial, in order to achieve efficiency as a function of itself. In both cases, it is believed that each user thinks not only in their own situation of use, but also in the process as a whole, coming to a result partially predicted as a concrete consequence.

The question that calls for great attention to User Experience researchers refers to the Popperian prerogative that you can only identify false experiences, and only when they are arranged in a favorable manner to refutation. As Bamford (2002) postulates, the logic of conjecture / analysis, instead of the consolidated analysis / synthesis, could contribute in the processes of Design in the extent that it puts in vogue the hypothesis under which its pretended to be analyzed. If in one hand it was possible to find, as referred to Philosophy of Design, an approach that is at the same time critical and progressive, on the other hand, one can not identify a coherent position on the User Experience methods - dealing with philosophical proposals that do not systematized, that merely resume ancient empirical techniques that only make evident the predominant scientism and pragmatism in Design.

It is true that the Design field, that to this point was guided by objective and concise criteria, is now being seriously questioned. If on one hand it highlights a fertile ground for new research, on the other there is a risk of no significant evolution in case our production happens in a unilateral way, without taking into account different points of view. The universal claim of user-faced models - which in fact are very effective in their contexts - manifested it self in statements that are intended to be self sufficient in any situation. Now, if we take Design in a too orderly and streamlined way, we can ignore our eternal user and subject matter: the human being. We believe this is the great role of Philosophy of Design on User Experience, regardless of what epistemological position is adopted: to examine and problematize, under a philosophical dignity, the methods used - leaving aside, even if provisionally, the so desired project solutions

Our considerations and questions could not take us any further than this. It would be necessary now raise other questions such as: User Experience must focus on the users intention or their uncertainties? Would the User Experience methods be a more effective mean to the Design than Design it self? A structured project focusing exclusively in User Experience, where the user guides the whole process, can still be considered a project of Design? How can we separate and measure the importance of each step in a User Experience project? This is what we would like to discuss in a future article

References

1. Baljon, C.J.: History of history and canons of design. Design Studies 23, 333–343 (2002)
2. Bamford, G.: From analysis/synthesis to conjecture/analysis: a review of Karl Popper's influence on design methodology in architecture. Design Studies 23, 245–261 (2002)
3. Bestelius, I., Doevendans, K.: Planning, design and the post-modernity of cities. Design Studies 23, 233–244 (2002)
4. Bevan, N.: What is the difference between the purpose of usability and user experience methods? In: Proceedings of the Workshop UXEM 2009 (INTERACT 2009). ACM Press, Uppsala (2009)
5. Broadbent, G.: Design and Theory Building. In: Cross, N. (ed.) Developments in Design Methodology. John Wiley, UK (1984)
6. Bucciarelli, L.L.: Between thought and object in engineering design. Design Studies 23, 219–231 (2002)
7. Calvera, A.: Treinando pesquisadores para o Design: algumas considerações e muitas preocupações acadêmicas. In: Design em Foco, vol. III(001), pp. 97-120. EDUNEB, janeira-junho, Salvador (2006)
8. Coyne, R.D.: Objectivity and the design process environment and planning. Planning and Design 19, 361–371 (1990)
9. Coyne, R.D., Snodgrass, A.: Is designing mysterious? Challenging the dual knowledge thesis. Design Studies 12(3), 124–131 (1991)
10. Coyne, R., Park, H., Wiszniewski, D.: Design devices: digital drawing and the pursuit of difference. Design Studies 23, 263–286 (2002)
11. Cross, N. (ed.): Developments in Design Methodology. John Wiley, UK (1984)
12. Cross, N.: Science and design methodology: a review. Research in Engineering Design 5, 63–69 (1993)
13. Cross, N.: Editorial: Forty years of design research. Design Studies 28, 1–4 (2007)
14. Daley, J.: Design creativity and the understanding of objects. Design Studies 3(3), 133–137 (1982)
15. Dilnot, C.: Design as a socially significant activity: an introduction. Design Studies 3(3), 139–146 (1982)
16. Dixon, J.R.: On a research methodology towards a scientific theory of design. In: Newsome, S.L., Spillers, W.R., Finger, S. (eds.) Design Theory 1988, Springer-Verlag, Berlin (1988)
17. Eder, W.E.: Report on workshop W3. In: Hubka, V., Eder, W.E. (eds.) Schriftenreihe WDK 7 Results of ICED 1981, Rome, Heurista, Zurich (1981)
18. Fontoura, A. M.: As manifestações pós-modernistas no desenho industrial e suas repercussões no ensino do projeto de produto. Dissertação de Mestrado em Educação – Pedagogia Universitária. PUC-PR, Curitiba (June 1997)

19. Galle, P.: Design as intentional action: a conceptual analysis. Design Studies 20(1), 57–82 (1999)
20. Houkes, W., Vermaas, P.E., Dorst, K., De Vries, M.J.: Design and use as plans: an action-theoretical account. Design Studies, 23, 303–320 (2002)
21. Kroes, P.: Design methodology and the nature of technical artefacts. Design Studies 23, 287–302 (2002)
22. Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J.: Understanding, scoping and defining user experience: a survey approach. In: Proceedings of the Human Factors and Computing Systems 2009. ACM Press, Boston (2009)
23. Liddament, T.: The computationalist paradigm in design research. Design Studies 20(1), 41–56 (1999)
24. Love, T.: Environmental and Ethical Factors in Engineering Design Theory: a Post positivist Approach. Praxis Education, Perth, Western Australia (1998)
25. Love, T.: Philosophy of Design: A Meta-theoretical Structure for Design Theory. Design Studies 21, 293–313 (2000)
26. Love, T.: Educating those involved in changing human futures: a more coherent programme for design education. In: Swann, C., Young, E. (eds.) Re-inventing Design Eucation in the University. School of Design, pp. 242–248. Curtin University of Technology, Perth (2001)
27. Love, T.: Constructing a coherent cross-disciplinary body of theory about designing and designs: some philosophical issues. Design Studies, 23, 345–361 (2002)
28. Margolin, V.: Design history or design studies: subject matter and methods. Design Studies 13(2), 104–116 (1992)
29. Oxman, R.: Educating the designerly thinker. Design Studies 20(2), 105–122 (1999)
30. Phillips, D.C., Burbules, N.C.: Postpositivism and Educational Research. Rowman & Littlefield Publishers, Lanham & Boulder (2000)
31. Sargent, P.: Design science or nonscience. Design Studies 15(4), 389–402 (1994)
32. Schön, D.: Towards a marriage of artistry and applied science in the architectural design studio. JAE 41(4), 4–10 (1988)
33. Thomas, J.C., Carroll, J.M.: The psychological study of design. Design Studies 1(1), 5–11 (1979)
34. Trott, E.: Permanence, change and standards of excellence in design. Design Studies 23, 321–331 (2002)
35. Zammito, J.H.: A Nice Derangement of Epistemes. Post-positivism in the study of Science from Quine to Latour. The University of Chicago Press, Chicago & London (2004)

Using Interaction Patterns in Heuristic Evaluation

Federico Botella[1], Jose A. Gallud[1], and Ricardo Tesoreiro[2]

[1] Operations Research Center University Institute, Miguel Hernandez University of Elche, Avda de la Universidad, s/n, Elche (Alicante), Spain
{federico,jgallud}@umh.es
[2] University of Castilla-La Mancha, Campus Universitario s/n, Albacete, Spain
ricardo.tesoriero@uclm.es

Abstract. Heuristics give some general principles or reflections that should be taken into account when an interface is being developed. Heuristic evaluation should end with a set of recommendations or advices directed to the responsible of the application or web site. Interaction patterns have reached a certain level of maturity. Designers and developers are offered different pattern catalogues that help them to design usable interfaces and better interactions. In this paper, a method to use interaction patterns in heuristic evaluation is proposed. The idea is to establish a correspondence between each heuristic and one or more interaction patterns. After presenting this correspondence, the paper illustrates the concepts by showing a real case.

Keywords: Usability, interaction patterns, heuristic evaluation.

1 Introduction

Nowadays, usability has becoming an important requirement in software projects. The traditional view of usability by which is something interesting or important, only when all the application has been developed and there is no more to be done, is obsolete. One of the most useful methods for evaluating the usability of applications and web sites, both for users and developers, is by making use of heuristics.

Heuristics give you some general principles or reflections that should be taken into account when the application interface is being developed. There are some well-known set of principles, as those defined by Nielsen [11], Shneiderman [17] and Norman [14]. Additionally, designers and developers can make use of other guidelines o good practices in order to assure to their applications the highest level of usability as possible.

These principles are offered as general sentences that have to be determined in each case. For example, one of these principles established that “the state should be always visible” that can be accomplished in different ways. This diversity can also introduce complexity in the process of usability evaluation.

Another problem is that the use of heuristics supposes the evaluator exhibits a rich experience in what usability means and how to check each principle in a particular case (form or web page). Sometimes is not possible to count on an expert for performing the evaluation. In these cases, the method proposed in this paper can help to do this kind of evaluation.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 23–32, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Heuristic evaluation should end with a set of recommendations or advices directed to the responsible of the application or web site. This important task requires expertise as the usability evaluation includes, not only negative aspects but also positive actions to improve the user interface. The idea described in this paper can also help in the task of proposing improvements.

On the other hand, interaction patterns have reached a certain level of maturity. Designers and developers are offered different pattern catalogues that help them to design usable interfaces and better interactions. Some of them are specific for web applications and other for all kind of applications.

In this paper, a method to use interaction patterns in heuristic evaluation is proposed. The idea is to establish a correspondence between each heuristic and one or more interaction patterns. After presenting this correspondence, the paper illustrates the concepts by showing a real case. This example of application also contributes to improve the method itself.

The paper is organized in the following sections. Section 1 covers the introduction and motivation of this research. Section 2 describes related work about heuristic evaluation techniques and methods. Section 3 offers a brief introduction about interaction patterns. Section 4 contains the proposal of heuristic evaluation using interaction patterns. A case study illustrates the proposal in section 5. Finally conclusions and future work is presented in section 6.

2 Heuristic Evaluation Techniques

A heuristic evaluation is a method to discover usability problems in user interfaces. In concrete, several evaluators examine the interface and judge if the interface accomplishes several usability principles, so called the “heuristics”. Usability inspection can be defined as a set of methods based on having evaluators inspect or examine usability-related aspects of a user interface [10]. In this paper, Nielsen and Molich described heuristic evaluation as “an informal method of usability analysis where a number of evaluators are presented with an interface design and asked to comment on it”. They presented here nine usability heuristics, and later Nielsen refined it to his famous “Ten Usability Heuristics” [19], based on a factor analysis of 249 usability problems [12] (see Table 1).

Nielsen’s heuristics are the most used list by many professionals for product evaluation. But other authors also defined their own lists of heuristics, like Shneiderman’s eight golden rules of interface design [17], Gerhardt-Powals research-based guidelines [8] or Norman’s seven principles [14].

Evaluators can define their own list of heuristics for a specific environment when they have to evaluate a concrete task, instead of restrict to use one of multiple list of heuristics published [24].

Heuristic evaluation can be performed by a small number of evaluators but several studies have demonstrated that the evaluation is more effective when multiple evaluators are involved. In fact, Nielsen defined that five-six evaluators will be sufficient to discover around 80% of the main usability problems of an interface [11].

Table 1. Evolution of Usability Heuristics from its conception

Nine usability Heuristics (Nielsen and Molich, 1990)	**Ten Usability Heuristics (Nielsen, 1994)**
1. Simple and natural dialog 2. Speak the user's language 3. Minimize user memory load 4. Be consistent 5. Provide feedback 6. Provide clearly marked exits 7. Provide shortcuts 8. Good error messages 9. Prevent errors	1. Visibility of system status 2. Match between system and the real world 3. User control and freedom 4. Consistency and standards 5. Error prevention 6. Recognition rather than recall 7. Flexibility and efficiency of use 8. Aesthetic and minimalist design 9. Help users recognize, diagnose, and recover from errors 10. Help and documentation

In heuristics evaluation sessions, when only one evaluator inspects the interface he uses to find a small number of usability problems. If we want to find other major usability problems we have to ask evaluators to fulfill a questionnaire after each evaluation session, where they can catalogue each one of usability problems discovered and moreover they can weigh each problem. Nielsen [13] proposed a severity scale for usability problems based on three factors: frequency of occurrence, user impact and persistence of the problem (see Table 2). One can see from the textual descriptions that priority is incorporated into the scale, and Nielsen has also noted that market impact should be assessed.

Table 2. Severity Ratings for Usability Problems

Numeric Value	**Textual Description**
0	I don't agree that this is a problem at all
1	Cosmetic problem only: need not be fixed unless extra time is available on project
2	Minor usability problem: fixing this should be given low priority
3	Major usability problem: important to fix, so should be given high priority
4	Usability catastrophe: imperative to fix this before product can be released

Furthermore, the discount usability engineering method [11] uses four techniques: user and task observation, scenarios, simplified thinking aloud and heuristic evaluation. Nielsen [13] suggests that there are two major reasons for alternating between heuristic evaluation and user testing. First, heuristic evaluation does not require users, who can be hard to get, to find usability problems. And second, many researches have demonstrated that both methods can find different problems.

3 A Brief Review about Interaction Design Patterns

Patterns, in general sense, have been defined to transmit understandings of design problems, by means of catching the core of recurrent problems and their solutions in a packed form. Patterns describe the problem fully, give the motivation of the solution and provide a form to apply the solution. A pattern language can be defined as a structured method for describing good design practices within a field of expertise.

Patterns are different from other forms of catching design knowledge, as guidelines or heuristics, in three ways. First, patterns offer a solution to concrete problem rather than providing abstract suggestions. Second, patterns are creative, helping designers, architects and engineers to generate a new solution by presenting several examples of the same problem. And third, patterns are connected between them by a hierarchical structure, so designers can resolve both complex and simple problems. Patterns are not going to replace guidelines or heuristics; patterns are a good complement for them.

Design patterns were developed by Christopher Alexander, in collaboration with Sarah Ishikawa and Murray Silverstein, as a method to allow anyone to design and build at any scale [1]. They based on the idea that users know more about the buildings they need than any architect could, idea that has been exported to the design of websites [22].

Alexander assumed that patterns could entitle both architects and their clients by providing a common and comprehensible language for design. He and his colleagues developed 253 patterns for building and planning towns, communities, houses, gardens and rooms. They wanted to outline an entire language for design, since the effectiveness of patterns was to provide a good solution to common problem, and at the same time to see how this solution was agreeable between client and architect [18].

The design patterns of Alexander are composed of a unique name, a numerical ID, an overview of the context and what the solution is about, regularly in the form of a brief summary and one picture. This overview is followed by a detailed description of the problem, how to implement the solution, a justification why the solution is good and the context where the design pattern should be applied [1].

UI designers [3,6] and also web designers [2,22,23] have adopted too the concept of design patterns. Some authors have described a collection of software design patterns which are now widely used [7]. We can found a main difference between software design patterns and design patterns of Alexander: the former were developed by and for professionals of the software industry while the second were particularly designed for creating good designs by non-professionals.

Design patterns ha0076e reached their greatest success in the area of software design by the success of the Gang of Four book Design Patterns [7], and also by the extensive practice of their pattern names in software development community. Moreover, design patterns have also reached to the area of human-computer interaction (HCI). The first HCI patterns can be considered at the Pattern Languages of Programming (PLoP) Conference [15] whereas pattern workshops began emerging at the Computer-Human Interaction (CHI) Conference [4]. Since then many pattern libraries have been published [22,23,20,25,6]. In figure 1 we can see an example of a typical interaction design pattern extracted from the web site of Van Welie [16]. The implementation of the interaction design patterns made by van Welie is very similar to the internal structure of a pattern form developed by Alexander; only the names of the attributes are slight different.

There are other fields of application of interaction design patterns, like e-learning, related to Computer Human Interaction. Van Diggelen and Overdijk developed design patterns for networked learning in the classroom and evaluated different solutions to several problems as design patterns [21]. Kohls and Uttecht presented a case study on the mining, writing and application of patterns for interactive educational graphics; it focuses on pattern mining and describes how to derive patterns from experience and analysis [9].

Dearden and Finlay [5] presented a critical review of patterns and pattern languages in human-computer interaction and introduced the concept of interaction design pattern to define design patterns in the HCI area. Thus, interaction design patterns are noticeably different from software design patterns which are focused on the source code and software structures rather than interface design.

Interaction design patterns have been developed for professionals and non-profesionals, as the original design patterns of Alexander. To create successful interactive systems, user interface designers need to cooperate with developers and application domain experts in an interdisciplinary team. These groups, however, usually miss a common terminology to exchange ideas, opinions, and values. Borchers presented an approach that uses pattern languages to capture this knowledge in software development, HCI, and the application domain. A formal, domain-independent definition of design patterns allows for computer support without sacrificing readability, and pattern use is integrated into the usability engineering life cycle [3].

Finally, we can consider that design patterns are basically a method for structuring knowledge and not a method for finding new solutions to problems. Solutions described in design patterns are not necessarily be new or original but should be demonstrated to work in practice. Therefore, design patterns are not developed from theory but identified as invariant aspects of solutions that emerge as best practices [5].

4 Using Interaction Patterns in Heuristic Evaluation

In this paper we propose a methodology to introduce the concept of interaction design patterns in the process of evaluation of a website by heuristics. To conduct a heuristic evaluation, first we have to define the set of evaluators who are going to examine the interface and their level of expertise; next we have to define a concrete task to evaluate and finally we have to provide a set of indicators (heuristics) to compound the questionnaire for evaluating the website. At this stage, we wanted to establish a correspondence between each one of the selected heuristics and one or more interaction design patterns in order to facilitate same directives to designers that helped them to overcome each usability problem.

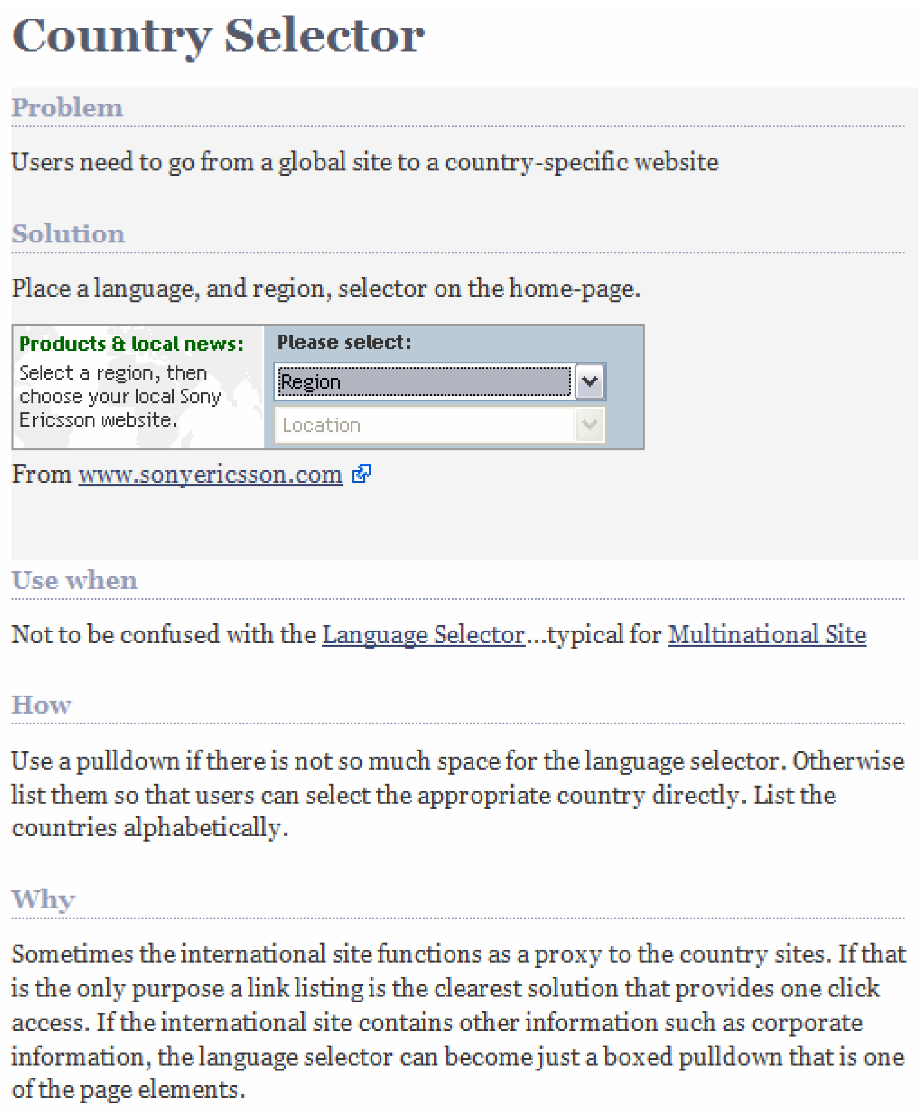


Fig. 1. Capture of a pattern of the web site of Van Welie (welie.com)

The set of indicators was selected as a set of heuristics grouped under the Ten Usability Heuristics of Nielsen [19]. On the other hand, we selected as a set of interaction design patterns a subset of the patterns defined by van Welie [16]. This set of patterns is a well-accepted pattern library like other patterns libraries as Yahoo Design Patterns Library [25] or Designing Patterns Library of Tidwell [20]. We selected Welie's Library by his simplicity, easy of use and similitude with the previous patterns designed by Alexander.

The interaction desing patterns defined by van Welie are grouped in three main sets: user needs, application needs and context and design. The first set, user needs, is composed of nine subsets: navigate around, searching, shoping, basic interactions, dealing with data, making choices, giving input, personalizing and miscellaneous. The second set, application needs, is divided into three subsets: drawing attention, feedback and simplifying interaction. And the third set, context of design, is divided into three subsets: site types, experiences and page types. Other patterns libraries could be more complete than van Welie, but this library is structured hierarchically what facilitates the process of mapping between the set of heuristics and a subset of interaction patterns.

For each one of the indicators of the set of a group of heuristics, we assigned at least one interaction pattern based on the context of use of the pattern and the target of the heuristic. For instance, the interaction patterns feedback and experiences were assigned to the heuristic “error prevention”. In Figure 2 we can see the relationship that we define between each heuristic and each one of the group of interaction design patterns. The relationship was established based on the relevance of each one of the interaction patterns contained in each one of the group of patterns that we consider are more relevant to solve the possible usability problems derived from each one of the heuristics.

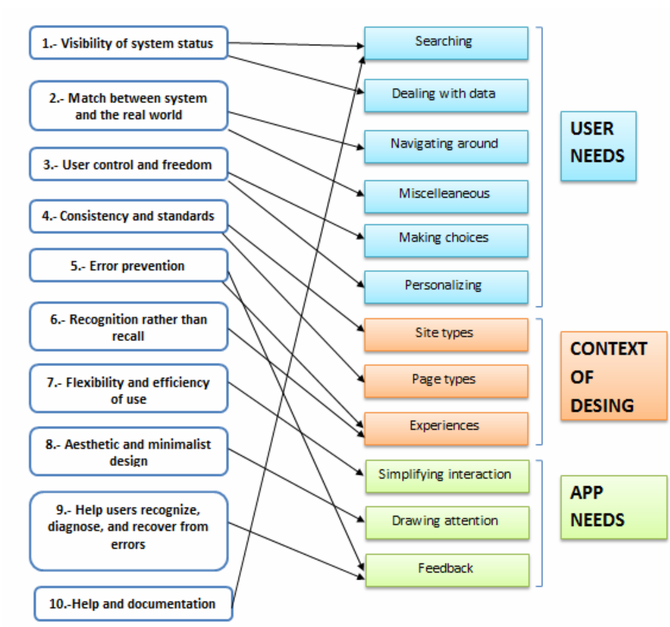


Fig. 2. Correlation between heuristics and interaction design patterns

The objective was to offer designers and developers a starting point to solve each one of usability problems detected in each heuristics when the evaluators diagnose the web site.

Finally we have defined a repository of evaluations with the aim to store all the evaluations performed by evaluators grouped by the type of website they evaluated (e-learning, e-commerce, intranet, etc.), the kind of task they evaluated (navigation, finding information, fulfilling a questionnaire), the usability problems detected and the suitability of the interaction pattern proposed to solve those usability problems. In case that the evaluator considered that other interaction pattern was more suitable to solve the usability problem detected, he can store the alternative pattern proposed. In this manner, after a number of heuristics evaluations we will get a valuable source of information for refining our correlation.

After each heuristic evaluation a report is issued indicating not only the usability problems detected but also a set of interaction patterns proposed to enrich the website evaluated. This will facilitate to designers and developers to solve the usability problems detected as they will receive a directed solution.

5 A Case Study

This section describes how the proposed method has been applied to a real Web site. It was not necessary to go far away to find a motivating and promising case study. This is why we decided to evaluate the web site of our university.

The goal was to improve the user experience of the whole community (students, professors and administrative staff) using the web site. This required performing the evaluation from different perspectives.

Fig. 3. The user has selected an option in the main menu but is difficult to see what

Figure 3 shows the home page of the site under study with a usability issue. As an example, the figure shows a submenu that has been dropped down by the user's selection in the main menu. It is not possible to determine whether the menu selection was “Investigacion” or was “Internacional” menu.

When heuristic evaluation was applied, we followed the Nielsen’s recommendation and we went through the interface twice. The first pass was intended to get a feel for the flow of the interaction and the general scope of the system. The second pass then allowed us to focus on specific elements testing how they fit the users’ need.

A number of usability issues were found. Some of them are listed following:

1. When the user navigates through the Web, (s)he does not know where (s)he is
2. The title of the different sections appear in no prominent places
3. It is difficult for the user to come back to the home page
4. The different sections do not have the same look and feel
5. The main menu does not give good feedback about the selected option

The usability report can conclude with a complete list of usability issues that must be corrected in the new version of the web site (We would like to remark that this homepage of the web site of our university was designed ten years ago, and precisely now we are involved in the process of opening a new web site, fully redesigned).

It is not necessary to define how to solve these problems; there are a lot of ways to fix each usability problem. The proposed method suggests that the usability expert takes the correlation we showed in the Figure 2 and try to find an interaction pattern that can be used to fix the problem

In the case of the former usability issues, the proposed solutions can be as follows:

1. Lost users in our web page can be fixed using the breadcrumbs pattern
2. The problem of no prominent information about sections can be solved by using the Header pattern
3. The problem of coming back to the home page can be fixed using the Home Link pattern
4. The problem of the different look and feel can be solved by using templates. This a more general problem that is fixed by a general rule
5. There are different possible solutions for the main menu. In the case of this web page we suggested to use the Double Tab Navigation pattern.

This only an example to illustrate how to use the pattern-heuristic correlation to provide a list of solutions that can be applied to different usability issues.

6 Conclusion and Future Work

In this paper we have proposed a new methodology to introduce the concept of interaction patterns into the process of heuristic evaluation.

As all usability evaluation should culminate with a report where all usability problems will be listed, introducing interaction design patterns will help to designers and to developers to better understand the usability problem detected and how to solve it directly as they receive a starting point to overcome the issue.

At the present, we are developing an environment to help evaluators to generate the usability report after a heuristic evaluation. The more evaluations done with the tool, the more interaction patterns will be possible to offer to new reports as the system will be able to select weighted interaction patterns that previous evaluators selected.

References

1. Alexander, C., Ishikawa, S., Silverstein, M.: A pattern language: Towns, Buildings, Construction. Oxford University Press, New York (1977)
2. Bayle, E., Bellamy, R., Casaday, G., Erickson, T., Fincher, S., Grinter, B., Gross, B., Lehder, D., Marmolin, H., Moore, B., Potts, C., Skousen, G., Thomas, J.: Putting it all together: Towards a pattern language for interaction design. In: A CHI 1997 workshop, pp. 7–23. ACM SIGCHI Bulletin, New York (1998)
3. Borchers, J.O.: A pattern approach to interaction design. AI & Society 15, 359–376 (2001)
4. Computer Human Interaction Conference, `http://www.chi2011.org`
5. Dearden, A., Finlay, J.: Pattern Languages in HCI: A critical review. Human-Computer Interaction 21(1), 49–102 (2006)
6. Erickson, T.: The Interaction Design Patterns Page, `http://www.pliant.org/personal/Tom_Erickson/InteractionPatterns.html`
7. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design patterns: Elements of reusable object-oriented software. Addison-Wesley Longman Publishing Co., Inc., Boston (1995)
8. Gerhardt-Powals, J.: Cognitive engineering principles for enhancing human-computer performance. International Journal of Human-Computer Interaction 8, 189–211 (1996)
9. Kohls, C., Uttecht, J.: Lessons learnt in mining and writing design patterns for educational interactive graphics. Computers in Human Behavior 25(5), 1040–1055 (2009)
10. Nielsen, J., Molich, R.: Heuristic Evaluation of User Interfaces. In: Proc. ACM CHI 1990, pp. 249–256 (1990)
11. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
12. Nielsen, J.: Enhancing the explanatory power of usability heuristics. In: Proceedings ACM CHI 1994 Conference, Boston, pp. 152–158 (1994)
13. Nielsen, J.: Usability Inspection Methods. John Wiley & Sons, New York (1994)
14. Norman, D.A.: The Design of Everyday Things. MIT Press, Cambridge (1998)
15. Pattern Languages of Programs (PLOP) Conferences, `http://www.hillside.net/index.php/conferences/plop`
16. Patterns in interaction design, `http://www.welie.com`
17. Shneiderman, B.: Designing the User Interface, 3rd edn. Addison-Wesley, Reading (1998)
18. Summary of the pattern language, `http://www.patternlanguage.com/leveltwo/patternsframegreen.htm?/leveltwo/./apl/twopanelnlb.htm`
19. Ten Usability Heuristics, `http://www.useit.com/papers/heuristic/heuristic_list.html`
20. Tidwell, J.: Designing interfaces: Patterns for effective interaction design. O'Reilly, Sebastopol (2005)
21. Van Diggelen, W., Overdijk, M.: Grounded design: Design patterns as the link between theory and practice. Computers in Human Behavior 25(5), 1056–1066 (2009)
22. Van Duyne, D., Landay, J., Hong, J.: The Design of Sites: Patterns, Principles, and Processes for Crafting a Customer-centered Web Experience. Addison-Wesley, Boston (2002)
23. VanWelie, M., Van der Veer, G.: Pattern languages in interaction design: Structure and organization. In: Proceedings of IFIP INTERACT 2003, pp. 1–5 (2003)
24. Xerox Heuristic Evaluation Checklist, `http://www.stcsig.org/usability/topics/articles/he-checklist.html`
25. Yahoo design pattern library, `http://developer.yahoo.com/ypatterns`

Making the Design Process More Usable: Aligning Design with User Performance

Roberto K. Champney, Christina Kokini, and Kay M. Stanney

Design Interactive Inc., 1221 E. Broadway Suite 110,
Oviedo Fl 32765, United States
{Roberto.Champney,Christina.Kokini,
Kay.Stanney}@designinteractive.netx

Abstract. One key drawback when evaluating usability return on investment (ROI) is that the assessment criteria are often subjective, making it difficult for members of a development team to buy in to the need to support usability-derived redesign recommendations. It is thus necessary to convey to the development team the importance of design for usability in a format that is universally understandable. The use of measurable usability requirements to assess usability ROI was found to be an effective approach to align design with operational performance and at the same time justify the need for redesign to the development team. This approach should result in better development team cohesion, as well as superior end product performance, which captures and supports the needs of end users and other stakeholders alike. In the current effort, this alignment process is described, and the utility of the approach is demonstrated by its application in a field case study of the successful design of a software application.

Keywords: Usability, Return on Investment, User Experience, Design.

1 Introduction

The importance of usability, in particular the importance of including user-centered evaluations from the beginning of the design lifecycle, has been common knowledge to human factors professionals for years [1]. However, other stakeholders, including members of the development team, often see such "soft" assessments as a hassle or even a luxury. While a number of studies have documented the value of usability, in practice [2] it is often necessary to continuously demonstrate the return on investment (ROI) such activities have to the overall development effort. One key drawback when evaluating usability ROI is that the assessment criteria are often subjective, which makes it difficult for members of a development team to buy into the need to support usability-derived redesign recommendations. It is thus necessary to convey to the development team the importance of design for usability in a format that is universally understandable. In the current effort, this process is described, and the utility of the approach is demonstrated by its application in a field case study of the successful design of a software application.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 33–42, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Background

Design and development is a multi-disciplinary effort that must take into consideration the input from the multiple disciplines involved. This often involves an iterative approach where designs are evaluated by usability practitioners alone or with end-users and result in the identification of usability shortcomings (i.e., issues). These evaluations are formative in nature [3] and thus are diagnostic, iterative, and focus on identifying usability problems that require resolution before a design is released for widespread use by its intended user population [4,5]. This is in contrast to summative evaluations which focus on quantifying usability performance. In ideal conditions, these findings would be readily transformed into design recommendations that would be instantiated into the system under review. Yet given the multi-disciplinary nature of development teams, the identified usability issues must often be carefully communicated to the other members of the team before they can be addressed so that they can understand them and also because usability issues must compete with a broad range of other priorities (e.g., time and budget, project goals, technical challenges). This creates a challenge since such reports must convey the importance of the findings in a way that other team members can understand and relate to such that they will include usability issues as a top priority.. In such an environment it becomes imperative that usability practitioners develop means to convey the criticality of their findings and align such findings to tangible metrics of performance. The National Institute of Standards and Technology (NIST) recognizes this as described in their Industry USability Reporting (IUSR) Formative Project [6] where they state "…With many more techniques available for this type of work, practitioners need clear definitions goals and guidelines to use in planning and conducting formative evaluations, especially in communicating/reporting the work (and its value) to colleagues and customers." (pg. 3). This is particularly important because often times it is the software development and management components of the team who assess the validity and criticality of usability findings, and not the usability practitioner [7].

For usability practitioners, this challenge is at the heart of their value proposition and thus several approaches have been suggested in the past to help address it. Formatting guidelines have been a popular approach which suggests that usability reports should be formatted for their intended audience. More specifically, reporting guidelines prescribe what elements or characteristics a report should include (e.g., drafting short-reports, an executive summary, usability issue severity classification, frequency of issues, positive findings, professional layout) [8]. Other strategies are more involved and prescribe the indoctrination of other stakeholders (e.g., the development team) into the value of the usability process in order to develop empathy for usability issues [7, 9]. These approaches involve the participation of the development team during user interviews, field visits, user testing, etc. or in exercises involving the prioritization of observed usability issues in order to develop a shared mental model of users' needs. While such approaches are of great value towards increasing the effectiveness of communicating usability findings and promoting the redesign of a system, they may be too time- and resource-intensive for the development team. These approaches also lack the ability to measure and convey the system's level of success from a usability perspective and are not prescriptive towards which issues affect project goals set by the entire development team. In addition, such approaches still do not provide sufficient insight to decision makers in the development team who may not be intimately familiar with the development work taking place, but who have direct input into the direction of an effort. In order to be more effective, usability evaluation reports need to incorporate objective, measureable operational objectives that can be tied directly back to project goals.

3 Proposed Approach

This work presents an alternative approach to supplement the reporting guidelines proposed by earlier work. The proposed approach incorporates the definition and conveyance of usability objectives that are measureable and aligned with project goals. The usability alignment process follows the traditional development lifecycle, including best practices for user-centered design, yet there are key differences to supplement and emphasize the focus on aligning usability needs to design elements, ensuring that project goals are met.

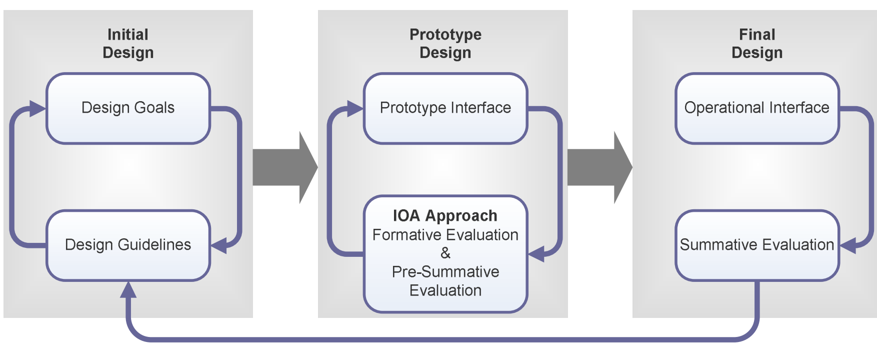


Fig. 1. Iterative Operational Assessment (IOA) Approach

What is proposed herein is an Iterative Operational Assessment (IOA) approach, as illustrated in Figure 1, which takes place iteratively during the development of a tool or system before it culminates in a final operational design. Traditionally formative usability evaluations take place once a prototype design has been developed, and summative usability evaluations don't take place until a final design is ready [10]. The IOA approach similarly follows this method, yet it incorporates "pre-summative" evaluations at the prototype stage as illustrated in Figure 1. The term "pre-summative" is coined here in reference to the assessment of human performance to evaluate the capabilities of the system at the prototype stage. Such an approach allows the design to be measured against an objective target (i.e., an operational objective linked to the project's goals) iteratively such that deviations from a desired outcome can be addressed via redesign. In order to derive value from these pre-summative evaluations, the measures to be assessed must be carefully thought through and agreed upon by the development team. This process is described next and is followed by a case study describing how this approach has been successfully applied in the field.

4 The Iterative Operational Assessment (IOA) Approach

The IOA approach is composed of four steps, as illustrated in Figure 2, where the outcome of one step is fed into the next as described below.

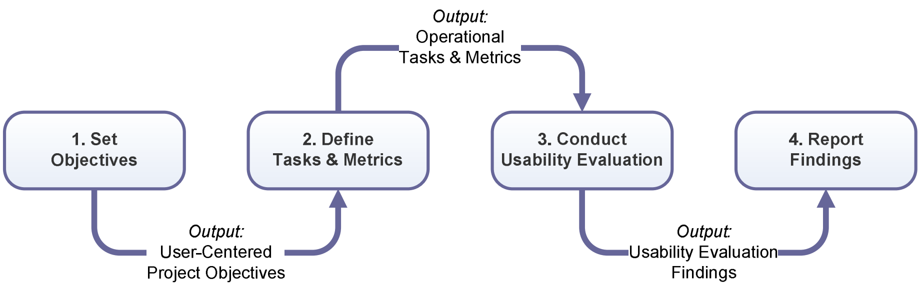


Fig. 2 The four steps of the IOA Approach

4.1 Set Objectives

As with traditional approaches, the process begins with the identification of relevant stakeholders, including decision makers, the development team and the usability practitioner as an advocate of the end-user. It is especially important to include decision makers that can identify the key project goals. The objective of this step is to provide alignment between project goals and user needs, and not to allow for decision makers to dictate design. One should always maintain user-centered vigilance throughout the process,yet it is necessary to obtain buy-in and insight from decision makers and other stakeholders. In this step (Table 1), the practitioner facilitates the identification of project goals via the application of user analysis methods with stakeholders (e.g., meetings, interviews, review of documentation). These goals are then decomposed into user-centered objectives (UC-Objectives) informed by the contextual field and task analyses that should have taken place earlier (as part of any usability engineering effort). Once the user-centered objectives have been defined, prioritized and agreed upon by stakeholders, it is time to move to the next step and define measures of expectation.

Lessons learned from the application of this method suggest that this approach ensures that decision makers are provided with data they understand and have agreed to, and thus eases any skepticism that may exist. Redesign recommendations

Table 1. Step 1 - Set Objectives

Process Component	Items
Inputs	Task and Domain Analysis, Project Scope and Objectives
Outputs	User-Centered Project Objectives
Participants	Decision Makers, Development Team, Usability Practitioner

associated with operational objectives (the UC-Objectives) set by the decision makers (as well as the rest of the stakeholders) are more likely to have more leverage during a more comprehensive prioritization of design requirements (i.e., design functionalities to support stakeholders', including end-users' tasks and goals) later in the process.

4.2 Define Tasks and Metrics

With the UC-Objectives (e.g., increase efficiency of field personnel) agreed upon on the prior step, it is now necessary to derive operational tasks and measures (Table 2) that would assess such objectives. In this step each UC-Objective is decomposed into related operational tasks (e.g., completing tasks X and Y) with the insight of the task and domain analysis work that should be part of any usability engineering effort. For each of these operational tasks, appropriate metrics (e.g., time to complete task) are developed such that performance by the system on these tasks can be objectively assessed.

Lessons learned suggest that these test tasks should be designed to achieve the same contextual objectives (e.g., select a training scenario that matches a prescribed training objective) as the baseline condition (i.e., prior to the design) in order to ensure objective comparisons.

Table 2. Step 2 – Define Tasks and Metric

Process Component	Items
Inputs	Task and Domain Analysis, User-Centered Project Objectives
Outputs	Operational Tasks and Metrics
Participants	Usability Practitioner

4.3 Usability Evaluation

The operational tasks and metrics defined in the prior step provide the guidance necessary to plan for a customary usability evaluation as. The key difference is that the tasks identified are now aligned not only with user needs, but also with project goals. Thus these contextually-relevant tasks become test tasks that can be used for iterative usability evaluations (Table 3). These evaluations can then assess system performance in the baseline condition (i.e., the way tasks are currently completed without the system being designed) and with each iteration of the system design. This does not imply that only user testing is conducted, in fact other evaluation methods should be utilized in conjunction, and the tasks identified in step 2 should be used to organize and prioritize the findings. A key difference in the way usability evaluations are conducted with the IOA approach is the "pre-summative" part of the evaluation. This is a stricter form of user testing than the traditional formative user testing since it requires a larger sample in order to provide a more reliable assessment of user performance. It is also not a true summative evaluation since the purpose is not to qualify the system to a desired level of performance, nor is the desired to achieve statistical significance, but to guide the design of the system such that it ultimately reaches a desired level of performance.

Lessons learned suggest that incorporating quantifiable usability requirements at this stage of the development lifecycle allows for the objective evaluation of the design's usability in terms of the key identified and agreed upon UC-objectives.

Table 3. Step 3 – Conduct Iterative Usability Evaluations

Process Component	Items
Inputs	Operational Tasks and Metrics
Outputs	Usability Evaluation Findings
Participants	Usability Practitioner, End-Users

4.4 Report Findings

The final step in the IOA approach is to report the findings such that the identified usability issues can be addressed. This is done in coordination with other formative usability evaluation methods which together with the quantifiable usability performance allow for definitive justification and prioritization of usability-derived changes to the design based on their impact. The reporting of the findings is done with input from available guidelines and tailored for their intended audience (e.g., the development team). The key contribution of the IOA approach is in how information is presented and prioritized. Given the alignment with the project goals,one can present the findings and redesign recommendations aligned with these goals in order to provide weight to their criticality. Further, objective comparison with the baseline and across design iterations provides a tangible indication to all members of the development team as to the capabilities of the design and reasons behind any redesign recommendations, thus providing a universal language to support the necessary communication in a multidisciplinary development team. Such alignment not only allows a development team to design and test iteratively for an objective performance target, but it also allows them to understand the progression of the design from a usability perspective.

5 Case Study

The IOA approach to formative usability evaluations has been applied in the field and the results of this real-life project are presented next as a proof-of-concept.

5.1 Background

During a recent project, the authors applied the IOA approach to a system being design to address the needs of the United States Marine Corps (USMC). The increases in training demands faced by today's Marines require that they take on more roles and master more knowledge, skills, and attitudes than ever before. Not only are more strenuous training requirements being levied, but at the same time the military is faced with the need to get troops trained with fewer resources. As a result, there is a need to increase the efficiency and effectiveness of learning to compress learning time and at the same time decrease instructor time requirements. The Deployable Virtual

Training Environment (DVTE) has been supplied as a potential solution to the challenge; it comprises a suite of laptop computers infused with a variety of training programs and simulation applications. Yet challenges still exist. To address some of the remaining challenges faced by the USMC, the Instruction Support Station (ISS) was designed and developed utilizing the IOA approach. Its purpose is to assist instructors using the DVTE system to conduct training. The application of this approach to the development of the ISS is discussed next.

5.2 Set Objectives

During the development effort of the ISS, the authors conducted a task analysis of the Marine training environment as is customary of human-computer interaction projects. This data was then utilized during multiple meetings with stakeholders, including decision makers and the multi-disciplinary development team to explicitly define key operational objectives for the project (i.e., the UC-Objectives). Two objectives are presented here: 1) Enhance instructor efficiency and 2) Enhance instructional effectiveness. These objectives did not only helpe define the theme of the effort but provided a means to prioritize the design effort into those design capabilities that would support these objectives.

5.3 Define Tasks and Metrics

Once the objectives had been defined and agreed upon, the authors, together with members of the development team, decomposed these objectives into operational tasks. Two examples are presented here: 1) Decrease the time to setup a training operation, and 2) Increase the ability to find suitable scenarios matching a desired training objective. With these operational tasks defined it was now possible to further define metrics to assess performance on these tasks, in this case: 1) time on task, and 2) whether or not a designated scenario was chosen given a particular training objective.

5.4 Usability Evaluation

The tasks and metrics defined were then integrated into the usability evaluation protocol that was used during the effort. As part of the user-centered design approach that took place, iterative design of low fidelity prototypes and multiple iterations of working prototypes were developed and evaluated. Both heuristic and empirical user-testing evaluations of the designs took place, allowing the authors to assess usability issues of the designs and address the issues via redesign recommendations. While the evaluations covered all of the aspects of the ISS design, the pre-summative user testing was focused on the key tasks defined and assessed via the metrics from the prior step. Together the findings of the evaluations were then integrated and reported.

5.5 Report Findings

In reporting the findings, guidance from best-practices was incorporated and integrated into redesign recommendations. The main difference from the traditional reporting approach was with the use of the pre-summative data. This data served to highlight current performance and to press upon the development team the need to implement the redesign recommendations pertaining to those key functionalities that affected the tasks representative of the project goals.

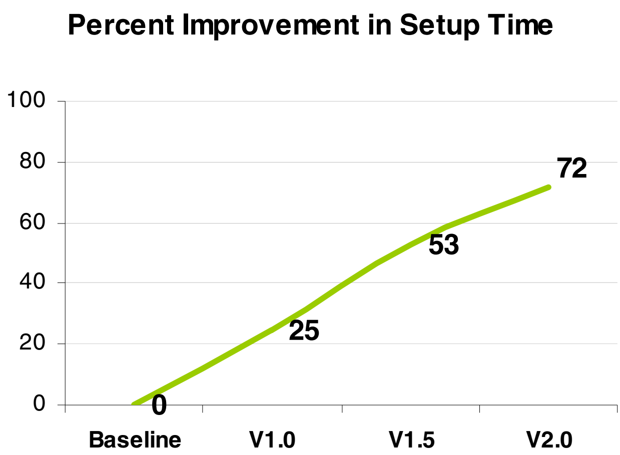


Fig. 2. Percent improvement in the setup time of a training operation from baseline

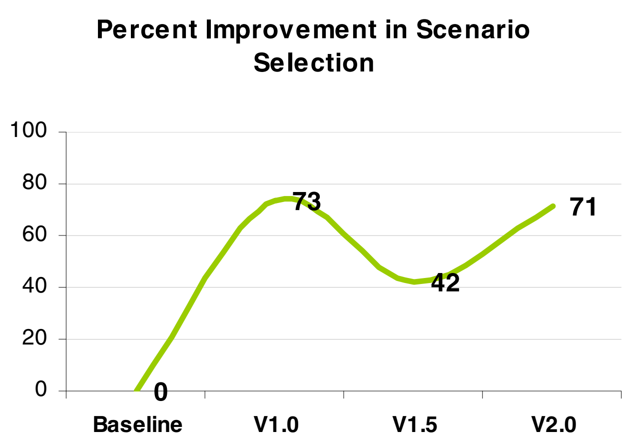


Fig. 3. Percent improvement in the selection of a designated scenario from baseline

6 Discussion

Through the iterative application of the IOA approach, the ISS design was guided to increase performance along those key tasks aligned with the project goal. In its current prototype version, the ISS has been able to achieve an improvement of 72% increase in efficiency of launching scenarios, and a 71% increase in the accuracy of selecting suitable scenarios for training. The authors found the IOA approach very powerful in quantifying the impact of the redesign recommendations. As shown in Figures 2 and 3, the iterative evaluations and the prioritization of redesign recommendations were necessary in order to maximize their impact on the project goals. One interesting observation is the potential interaction effects that may be observed during this process. During the redesign of the ISS V1.5, several features were integrated in order to increase the performance of one objective (see Figure 2) resulting in repercussions towards a second objective (see Figure 3). Yet though the iterative process of continuous improvement via redesign, performance was regained and expected to increase as the design further matures.

7 Conclusion

In sum, the use of quantifiable user-centered objectives to assess usability ROI was found to be an effective approach to align design with operational performance and at the same time justify to the development team the reasons for the need to redesign. This approach should result in better development team cohesion, as well as superior end-product performance, which captures and supports the needs of end-users and other stakeholders alike.

Acknowledgement. This material is based upon work supported in part by the Office of Naval Research (ONR) under contract N00014-08-C-0186. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views or the endorsement of ONR.

References

1. Mayhew, D.J.: The usability engineering lifecycle. Morgan Kaufmann Publishers, Inc., San Francisco (1999)
2. Bias, R.G., Mayhew, D.J.: Cost-justifying Usability. Morgan Kaufmann Publishers, Inc., San Francisco (2005)
3. Hewett, T.T.: The role of iterative evaluation in designing systems for usability. In: Harrison, M.D., Monk, A.F. (eds.) People & Computers: Designing for usability. Proceedings of the second conference of the BCS HCI specialist group, Cambridge University Press, Cambridge (1986); In: Booth, P.: An Introduction to Human-Computer Interaction, LEA, Hove, East Sussex, U.K. (1990)
4. Hartson, H.R., Andre, T.S., Williges, R.C.: Criteria for evaluating usability evaluation methods. International Journal of Human-Computer Interaction 13(4), 373–410 (2001)
5. Quesenbery, W.: Defining a summative usability test for voting systems: A report from the UPA 2004 Workshop on Voting and Usability (Online). Usability Professional's Association (2004), `http://www.usabilityprofessionals.org/upa_projects/voting_and_usability/documents/voting_summative_test.pdf` (retrieved January 20, 2008)

6. Theofanos, M., Quesenbery, W., Snyder, C., Dayton, D., Lewis, J.: Reporting on Formative Testing: A UPA 2005 Workshop Report. Usability Professionals Association (2005), http://www.usabilityprofessionals.org/usability_resources/conference/2005/formative/%20reporting-upa2005.pdf
7. Kneifel, A.A., Guerrero, C.: Using participatory inquiry in usability analysis to align a development team's mental model with its users' needs. In: Proceedings of the Society for Technical Communication 50th annual conference, Dallas, Texas, USA, May 18-21 (2003)
8. Dumas, J.S., Redish, J.C. A practical guide to usability testing, 2nd edn. Intellect, Portland, OR (1999)
9. Redish, J., Bias, R.G., Bailey, R., Molich, R., Dumas, J., Spool, J.M.: Usability in Practice: Formative usability evaluations – evolution and revolution. In: Proceedings of the Computer Human Interaction Conference, Minneapolis, Minnesota, USA, April 20-25 (2002)
10. Kies, J.K., Williges, R.C., Rosson, M.B.: Coordinating computer-supported cooperative work: A review of research issues and strategies. Journal of the American Society for Information. Science 49, 776–779 (1998)

Exploring the Learning Problems and Resources Usage of Undergraduate Industrial Design Students in Design Studio

Wenzhi Chen[1] and Hsien-Hui Tang[2]

[1] Department of Industrial Design, Chang Gung University, 330 Tao-Yuan, Taiwan, R.O.C.
[2] Department of Commercial and Industrial Design, National Taiwan University of Science and Technology, 106 Taipei, Taiwan, R.O.C.
wenzhi@mail.cgu.edu.tw, drhhtang@mail.ntust.edu.tw

Abstract. Design is a power weapon of modern companies. It is important to have excellent designers for the industry. The purpose of this study is to explore the learning problems, and the resources that students used to overcome problems in undergraduate industrial design studio courses. The survey with open type questionnaires was used to collect data. There were 189 undergraduate industrial design students from 3 universities participating in this study. The results demonstrated that the most difficulties design tasks included concept generation, design presentation, and design research. The learning resources used to solve the learning problems included 4 categories: people, object, method, and environment. This information can increase the understanding of the learning process of students, and provide the reference for teaching planning and the setting of the learning resources in design education.

Keywords: Industrial design, learning problems, learning resources, design studio.

1 Introduction

Design is a power weapon of the companies. It is important to have excellent designers for the industry. The essential feature of the design education is "learning by doing". Students learn design knowledge and skill through operating the real design problem. Based on the feature of the design education, the problems that students experienced and the resources used in the learning process may be different from those in other domains. But, there are few studies discussing the learning problems and learning resource in design education.

The purpose of this study is to explore the learning problems and the resources that students used to overcome problems in the undergraduate industrial design studio courses. The results can increase the understanding of the learning process of students, and provide the reference for teaching planning and the setting of the learning resources in design education.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 43–52, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Literature Review

2.1 Design Professional Learning

Learning is the act or process of developing skill or knowledge [1]. Design is generally considered to involve abductive reasoning, which addresses ill-defined problems and uses a construction process to solve them [2]. Design knowledge is difficult to understand or to describe. It must be understood with reference to the problem context. It cannot be delivered by the traditional lecture pedagogy. The most used way is "learning by doing", such as the apprentice system, so design learning emphasizes working with actual problems to acquire professional knowledge and techniques [3]. Students need to acquire advanced knowledge through working with design problems in the real world. In the learning process, a teacher demonstrates and leads students to engage with a real design problem. Students try to observe, to understand, and to grasp the various methods and techniques through the process, to cultivate observation and decision-making capabilities, to explore their own talents, and finally to develop their own procedure and style of design. Therefore, the design process can be considered as a social process in which the result is constructed using various kinds of knowledge [4].

The main and most important teaching method of design education is studio [3] that has been in use for almost 100 years [5], especially in architectural and industrial design. The main distinguishing feature of studio pedagogy is the learning of the procedure and methods of design and the accumulation of experience through the process of solving actual design problems. The emphasis is on the presentation of design concepts and ideas, the critique and communication involved in the design process, and the learning of advanced design knowledge through reflection on design problems. Students go through the procedure of design to solve real or simulated design problems [6-8]. Through the studio, students are exposed to a number of learning experiences, focusing on two key aspects. The first is learning how to design by engagement with a process of designing or a suite of possible design methodologies. The second is to reveal knowledge about concepts and/or situations through the act of designing [9].

Design studio is the heart of most industrial design and architecture curricula. There are several factors associated with a successful design studio. The "studio teaching project (http://www.studioteaching.org)" supported by the Australian Learning & Teaching Council proposed that there are several key qualities or characteristics that play important roles in a typically successful studio program: 1) People: lecturers, tutors, technicians, members of the professional communities and student peers; 2) Facilities and Resources: space, equipment, technologies and materials; 3) Projects: areas of study, tasks and problems to be solved, especially those related to industry/profession; and 4) Time: the proportion of course time provided for studio and hours of access to facilities. Attoe & Mugerauer [6] also mentioned about the factors associated with teaching excellence in design studios, including three considerations: 1) the teacher as self (aspects of the teacher's own life that contributes to good teaching); 2) personal style (the way the teacher behaves); and 3) course format and implementation.

2.2 Learning Problems and Resources in Design Studio

The design studio is the place to practice and integrate the knowledge and skills that learnt from the surrounding courses. There are some people involved with the design learning process, for example, classmate, instructors, technician, and other experts. It also needs environments and settings to support the design learning process, such as the personal working space, workshops, library, etc. Learning the process of design is similar to designing. Both events involve spending certain periods of time thinking through the process and attempting to create some interesting results. Learners often have to explore and discover their own paths to gain the knowledge and skills [10]. The students faced a number of problems as they worked through their design process.

Learning Problems

The issues of students' learning problems or difficulties are complex and dependent on a range of factors, including course organization and development, the subject or topic being taught, teaching style, and students' expectations [11, 12].

Design students faced some learning problems and difficulties when they explored in the design problems space. Yang, You, & Chen [13] investigated the difficulties faced by industrial design students and their career guidance needs. Focus group interview was used to collect the data. The participants included freshmen, seniors, graduates of senior high schools, and vocational high schools, with or without a design background. The problems students encountered were: 1) high learning pressure and frustration; 2) heavy workload depriving them of extra-curricular activities and leisure; 3) competition among classmates influencing peer relationships; 4) high costs of materials resulting in financial pressure; 5) feeling of uncertainty and worries about the future; and 6) insufficient interaction between faculty and students. Mawson [14] compared the workplace practice of six experienced designers and investigated their experience on teaching practicum in developing design skills with secondary school students. They found students faced 2 types of problems. The first were technical problems related to the skills needed to produce the product. The second type of problems related to the nature of the materials and the tools needed to do the job. The teachers also identified three basic problems with their experience. One was students' antipathy towards and ignorance of the design element in the technological process. Another was the lack of practical experience. A further problem was entrenched traditional views of technical education in the schools that were not congruent with the approach of the relatively new technology curriculum.

Learning Resources

Learning resources are defined as information, represented and stored in a variety of media and formats that assists student learning as defined by provincial or local curricula. This includes but is not limited to, materials in print, video, and software formats, as well as combinations of these formats intended for use by teachers and students [15].

Learning the process of design is similar to designing [10]. The design students need some resources to solve the learning problems and difficulties, as well as to solve the design problems with design resources and knowledge. Some design resources used to solve the design problems may also be the design learning resources for solving the learning problems.

Little research discusses about the design learning resources, even the study method. Brown, Doughty, Draper, Henderson & Mcateer [16] had been developed a learning resource questionnaire to gather information on the learning resources used by students. The resources may include not only lectures, tutorials and courseware, but books, handouts, notes and discussions with other students. They regarded this information is important to teaching staff in assessing and increasing the value of the resources to students by ensuring their effective integration into a course.

About the design learning resources, Chiu [17] has been investigated the students' knowledge sources and knowledge sharing in the design studio, the definition of the design knowledge is similar as the design learning resources in this study. The results demonstrated that the top four knowledge sources requested by juniors were books and magazines, studio-mates, schoolmates, and the Internet. The top four knowledge gain sources were books and magazines (40%), studio-mates (22%), schoolmates (12%), and the Internet (9%). The top four knowledge sources requested by seniors were books and magazines, the Internet, studio-mates, and auditing desk crits. In terms of percentage of knowledge gain, books and magazines were 25%, the Internet 23%, studio-mates 20%, and auditing desk crits 15%.

The research of You, Yang & Liao [18] explored the industrial design students' learning attitudes in Taiwan. Some results of the study also related with the learning resources, including: 1) while learning design, students aspire for teachers to share design experiences with them; 2) they would ask classmates or friends for help when encountering difficulties; 3) they devote much time and place emphasis on creative thinking and model-making during the design process; 4) their design concepts mostly come from their life experiences.

The main method for design education is studio, where students learn design knowledge through the design process with operating real design problems. Design students faced some learning problems and difficulties in design learning process, and tried using some learning resources to solve the problems just like searching and gathering the resources to solve the design problems. But little research discusses the learning problems and learning resources in design education domain. This paper tries to explore the learning problems and learning resources to increase the understanding of the learning process of design students.

3 Methods

A survey was conducted to explore the learning problems and the learning resources that used to solve the learning problems in industrial design studio. The main questions in this study include:

1. What are the most difficult tasks of the design process in design learning project?
2. What are the major problems that students experience in the each tasks of the design learning projects?
3. What kinds of the resources are used by students to overcome problems?

3.1 Subjects

The participants in this study were 189 undergraduate industrial design students from 3 different Universities in Taiwan. The subjects' attributes are shown as in Table 1.

Table 1. The subject numbers and percentages of school, gender, and year

	School			Gender		Year	
	U1	**U2**	**U3**	**Female**	**Male**	**2nd**	**3rd**
Frequency	42	127	20	116	73	88	101
Percentage (%)	22.2	76.2	10.0	61.4	38.6	46.6	53.4

3.2 Data Collection and Analysis

The data were collected using questionnaire survey. There are 3 categories questions in the questionnaire: 1) the learning problems that students experience in each design tasks; 2) the resources that students use to solve the problems in each design tasks; 3) the basic information of the subjects. For exploring the questions, open type questions were used in category 1 and 2 to collect the data. Both quantitative and qualitative data analyses were performed. First, descriptive statistics of the students responses were computed. Next, the content of problems and resources was coded. Finally, the content coding also was calculated and tested.

4 Results

4.1 Basic Information

The students spent on average 30.63 (SD = 19.48) hours per week for their design learning projects. They slept 5.47 (SD = 1.04) hours and spent 4.85 (SD = 2.50) hours surfing on Internet every day on average.

4.2 Learning Problems

Table 2 shows the percentage of each design task that students experienced with years, gender, and schools. The results show that the top 3 difficulties design tasks that students responded were *concept generation* (82.5%), *design presentation* (40.7%), and *design research* (34.4%). The chi-square tests were conducted to test of homogeneity of proportions between years, gender, and schools. The result indicates a significant difference between schools (X^2 = 16.361, df = 8, p = 0.037 < 0.05). The top 3 tasks that students experienced problems of U1 were *concept generation* (73.8%), *design presentation* (54.8%), and *design decision* (50.0%); of U2 were *concept generation* (85.8%), *design research* (37.8%), and *design presentation*

(36.2%); of U3 were *concept generation* (80%), *design research* (40%), and *design presentation* (40%). The U1 students seem to face more problems than U2 and U3, especially in *design decision* task. But U2 and U3 students seem to have more problems in *design research* and *design documentation*. However, no significant difference in gender (X^2 = 3.052, df = 4, p = 0.549 > 0.05) and year (X^2 = 2.676, df = 4, p = 0.613 > 0.05).

Table 2. The percentages of problems that students faced in terms of year, gender, and school; and the results of the chi-square tests

Resource	Year		Gender		School			Total
	2^{nd}	3rd	Female	Male	U1	U2	U3	
Design Research	34.1	34.7	32.8	37.0	21.4	37.8	40.0	34.4
Concept Generation	89.8	76.2	81.9	83.6	73.8	85.8	80.0	82.5
Design Decision	34.1	28.7	33.6	27.4	50.0	25.2	30.0	31.2
Design Presentation	39.8	41.6	40.5	41.1	54.8	36.2	40.0	40.7
Design Documentation	15.9	7.9	14.7	6.8	2.4	15.0	10.0	11.6
X^2		2.676		3.052			16.361	
df		4		4			8	
P Value		0.613		0.549			0.037*	

Note: percentage = number of responses / number of subjects.
* p < 0.05.

The learning problems students responded were coded and divided into three categories: 1) *personal* (89.6%); 2) *resources* (6.1%); and 3) *interaction and communication* (4.3%). The *personal* problems included capabilities, thinking, technique and skills, experiences, personality, knowledge, and others issues. The *resource* problems were money (cost), time, technical support, equipment, and related courses. The *interaction and communication* problems were occurred with instructors and peers. Table 3 presents the percentages of the problem categories in each design tasks.

The top 3 major problems of design research task were *personal* thinking (49.3%), capabilities (14.3%), and experiences (8.5%); of concept generation were *personal* thinking (53.7%), capabilities (10.1%), and experiences (9.4%); of design decision were *personal* capabilities (31.0%), thinking (23.1%), and *interaction and communication* with instructors (9.9%); of design presentation were *personal* capability (38.2%), technique and skills (30.3%), and personality (6.8%). The *personal* capabilities (60.2%), thinking (9.7%), and technique and skills (7.1%) were the top 3 problems in design documentation tasks.

Table 3. The percentages of the problem categories in each design tasks

Problem\Task	Design Research	Concept Generation	Design Decision	Design Presentation	Design Documentation
Personal	**92.1**	**92.5**	**83.3**	**88.0**	**91.7**
Capability	14.3	10.1	31.0	38.2	60.2
Thinking	49.3	53.7	23.1	3.1	9.7
Technique & Skill	10.2	7.8	8.0	30.3	7.1
Experience	8.5	9.4	6.8	3.3	2.0
Personality	2.0	1.8	4.4	6.8	2.0
Aesthetics	1.8	4.1	4.9	1.9	2.2
Knowledge	4.1	3.3	2.6	2.0	1.7
Other	2.3	2.5	3.2	3.1	7.1
Resource	**5.6**	**2.7**	**3.0**	**11.6**	**6.9**
Money (Cost)	0.7	0.5	1.8	4.4	0.6
Time	0.7	0.3	0.6	3.1	1.4
Technical Support	0.1	0.1	0.2	2.8	2.3
Equipment & Tool	0.3	0.3	0.0	2.3	0.9
Related Course	0.0	0.1	0.2	0.8	0.3
Other	4.0	1.9	0.3	1.3	2.5
Interaction & Communication	**2.3**	**4.8**	**13.7**	**0.3**	**1.4**
Instructor	2.1	3.0	9.9	0.3	1.2
Peer	0.1	1.9	4.4	0.0	0.2

Note: percentage = number of responses / total number of responses in task category.

4.3 Learning Resources

The learning resources that students responded were divided into 4 categories: the *people* (47.3%), *object* (30.5%), *method* (22.6%) and *environment* (3.2%). Table 4 shows the percentages of resource categories between years, gender, and schools. The chi-square test indicates a significant difference between year (X^2 = 13.189, df = 3, p = 0.004 < 0.01), and schools (X^2 = 223.417, df = 6, p = 0.000 < 0.001). It seems that the 2nd year students depend on the *people* resources more, while 3rd year students begin trying to use *object*, *method*, and *environment* resources to solve the learning problems more. The U1 seems to use the *object* and *method* resources more often than others schools, and U2 & U3 seems to depend on the *people* resource more. However, there is no significant difference in gender (X^2 = 4.111, df =36, p = 0.250 > 0.05).

Table 4. The percentages of learning resources categories in year, gender, school; and the results of the chi-square tests

Task	Year		Gender		School			Total
	2nd	3rd	Female	Male	U1	U2	U3	
People	48.9	46.0	46.7	48.3	40.5	50.8	43.9	47.3
Object	29.5	31.3	31.2	29.4	32.9	29.6	29.4	30.5
Method	21.9	23.3	22.6	22.7	32.2	18.5	22.0	22.6
Environment	2.7	3.5	2.8	3.7	3.8	2.5	5.5	3.2
X^2		13.189		4.111			223.417	
df		3		3			6	
P Value		0.004**		0.250			0.000***	

Note: percentage = number of responses / total number of responses in task category.
* p < 0.05, ** p < 0.01, *** p < 0.001.

Table 5 shows the resource categories used in each design task. The chi-square test indicates a significant difference between design tasks (X^2 = 241.371, df = 12, p = 0.000 < 0.001). The *people* resource was most used in each task, especially in design decision. Students used the *people* and *object* as main resources for solving the problems in most design tasks, but used *method* (25.0%) more than *object* (21.0%) resources in design decision task.

Table 5. The percentages of learning resources categories in each design task

Resources\Task	Design research	Concept generation	Design decision	Design presentation	Design docummentation
People	44.5	40.5	56.7	50.2	44.9
Object	35.1	34.0	21.0	25.9	22.8
Method	20.5	26.1	25.0	22.8	18.5
Environment	4.0	3.9	1.6	4.0	2.2

Note: percentage = number of responses / total number of responses in task category.

5 Concluding Remarks

The purpose of this study was to investigate the learning problems that students faced and the resources used for undergraduate industrial design students in studio courses. The results indicate that 1) the *concept generation* was the most difficulty design task since 82.5% students experienced the problems; 2) the main problems were related with students' *personal* issues; 3) the *people* resource was most used for solving the problems, especially in *design decision* task; 4) different schools had significant difference between their learning problems and resources usage. These findings are consistent with previous studies [13, 17, 18]. Several additional findings and reflections are elaborated below.

The Capabilities and Thinking Styles of Students

The research results indicated that students experience lots of problems related to their capabilities and thinking styles. There were 22.6% students using *methods* to solve the problems, and 47.3% using *people* resources. Design studio is the core course of the design education, so students need to implement the knowledge and techniques learned from the supporting courses. The curriculum design may influence the performance [19]. Are the supporting courses appropriately arranged to support the core course? Do students really learn the abilities? Do they know how to apply the abilities learnt? These questions will significantly related to the problems students will encounter.

Resources Setting and Supporting

The students responded they faced problems related to *resources* setting, such as money (cost), time, technical support, equipment, and *interaction* with instructors. Reviewing the answers found that some issues were related to students themselves, for example, the time management skill. However, there were some related to facilities and the technical support, such as the workshops and processing machines. Therefore, design learning is similar to apprentice system that students learn design through doing the design with instructors' guidance. The instructors' support is important [18, 20]. In this study, some students seem to have problems to interact with instructors. Some of them had different opinions with instructors, especially in design decision. It is important to guide students dealing with conflicts and making decisions.

Design professional learning is a complexity issue. This study is an exploratory research, and therefore the findings were not very conclusive. Future research will provide more details to provide insights regarding to learning problems and resources usage.

Acknowledgments. This research was partly funded by the National Science Council of Republic of China under grant numbers NSC 99-2410-H-182-028-MY2, and Chang Gung University under grant number UARPD380111.

References

1. Arsham, H.: Impact of the Internet on learning and teaching. USDLA Journal 16(3) (2002)
2. Zimring, C., Latch Craig, D.: Defining design between domains: an argument for design research? la caret. In: Eastman, C.M., McCracken, W.M., Newsletter, W.C. (eds.) Design knowing and learning: cognition in design education, pp. 125–146. Elsevier, Amsterdam (2001)
3. Schön, D.A.: Education the reflective practitioner. Jossey-Bass, London (1987)
4. Bucciarelli, L.L.: Design knowledge & learning: a socially mediated activity. In: Eastman, C.M., McCracken, W.M., Newsletter, W.C. (eds.) Design knowing and learning: cognition in design education, pp. 297–314 (2001)
5. Reimer, Y.J., Douglas, S.A.: Teaching HCI design with studio approach. Computer science education 13(3), 191–205 (2003)

6. Attoe, W., Mugerauer, R.: Excellent studio teaching in architecture. Studies in Higher Education 16(1), 41–50 (1991)
7. Budd, J., Vanka, S., Runton, A.: The ID-Online asynchronous learning network: a 'Virtual Studio' for interdisciplinary design collaboration. Digital Creativity 10(4), 205–214 (1999)
8. Forgber, U., Russell, P.: Interdisciplinary collaboration in the virtual design studio, Plymouth, England (1999)
9. Smith, D., Hedley, P., Molloy, M.: Design learning: a reflective model. Design Studies 30(1), 13–37 (2009)
10. Chang, T.-W., Huang, J.-H.: A pilot study of role-interplay in a web-based learning environment. Education Media International 39(1), 75–85 (2002)
11. Chang, P.-F., Hsiau, S.-S., Yeh, T.-L., Wu, J.-C.: The development and implementation of the technological creativity course: an interdisciplinary approach. In: International Conference of Engineering Education (ICEE), Taipei, Taiwan (2000)
12. Pirrie, A., Hamilton, S., Wilson, V.: Multidisciplinary education: some issues and concerns. Educational Research 41(3), 301–314 (1999)
13. Yang, M., You, M., Chen, F.: A study on the difficulties and career guidance needs of industrial design students: implications for design education. Journal of Design (TW) 10(2), 57–76 (2005)
14. Mawson, B.: Designers as teachers and learners: Transferring workplace design practice into educational settings. International Journal of Technology and Design Education 17(2), 163–177 (2007)
15. Learning Resources, `http://www.bced.gov.bc.ca/irp/appskill/asleares.htm#def`
16. Brown, M.I., Doughty, G.F., Draper, S.W., Henderson, F.P., Mcateer, E.: Measuring learning resource use. Computers & Education 27(2), 103–113 (1996)
17. Chiu, S.-H.: Students' knowledge sources and knowledge sharing in the design studio—an exploratory study. International Journal of Technology and Design Education 20(1), 27–42 (2010)
18. You, M., Yang, M.-Y., Liao, P.: Survey of industrial design students' learning attitudes. Journal of Design (TW) 12(2), 15–36 (2007)
19. Tempelman, E., Pilot, A.: Strengthening the link between theory and practice in teaching design engineering: an empirical study on a new approach. International Journal of Technology and Design Education (in press)
20. Studio Teaching Toolkit: What is Studio? Current Models, `http://www.studioteaching.org/index.php?page=current_models`

Towards Future Methods to Take into Account Cross-Cultural Differences in Design: An Example with the "Expert Community Staff" (ECS)

Jérôme Dinet, Robin Vivian, and Eric Brangier

Université Paul Verlaine – Metz. InterPsy-ETIC, EA 4432. User Experience Lab. Faculté des Sciences Humaines et Arts BP 30309 Île du Saulcy - F-57006 Metz, France
{jerome.dinet,robin.vivian,brangier}@univ-metz.fr

Abstract. Nowadays, ergonomics tries to imagine and to create new methods based on social situations to understand users' needs when these end-users are issued from different cultures. Because these needs are socially and culturally determined, new technology cannot be designed without understanding how it is embedded in its socio-cultural context. The aim of this paper is to present the Expert Community Staff (ECS) method, an innovative participatory method to enlarge real users' needs. First, we present the theoretical background of this method. Second, we present the methodological implications for the design. Finally, the advantages and disadvantages of this new participatory method are discussed.

Keywords: participatory method, users' needs, socio-constructivist design, community of users.

1 Introduction

This article proposes an innovative participatory method to take into account cross-cultural differences in design of new technology, to complement the traditional methods used in ergonomics.

We argue that it is essential to create innovative methods to include users issued from different cultures, with different histories, knowledge and with levels of perceptual, cognitive and social skills because new questions appear for the designers such as: "How to represent end-users who do not exist yet, because the technology has not been developed?" "What kinds of method exist to enlarge the representation of end-users who are scattered around the world?".

To answer these questions, most research in ergonomics has concentrated on a single dimension— either psychological or social—although a few researchers have considered both in a single study. The psychological approach focuses fundamentally upon attributes of the individual and examines psychological motivations for information behaviour that cut across several contexts or is independent of context [10, 27, 32]. The social approach studies the effects of the social context on information behaviour, regardless of an individual's psychological attributes [9, 30]. This article aims to consider an alternative to these two conventional approaches (either the psychological or the social) by adopting a constructivist perspective: the "Expert Community Staff" (ECS).

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 53–61, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Theoretical Background of the ECS Method

2.1 From Psychological and Social Approaches

As Allard, Levine & Tenopir [1] said, researchers in the domain of human behavior have traditionally elected to conduct their work within either the psychological or the social approach. The psychological approach focuses primarily on a study of psychological states and processes in relation to information behavior. Central to this approach is the concept of information needs [17, 19, 33]. Based on the studies conducted by Taylor [31], the first approximation of the concept of information needs was proposed [3, 12, 13].

A lot of studies employed these approaches to examine information behavior at various cognitive levels, such as stages in solving a problem [7, 21], the level of familiarity with a topic [23], the level of task complexity [8], and along individual cognitive factors [18]. Nevertheless, the majority of creators of theoretical frameworks and models in the psychological tradition would probably agree that additional factors, which are not necessarily psychological (such as the physical environment, socio-cultural background, and goals of a search), also affect information behavior. For instance, Dervin [12, 13] emphasized that information behavior is also determined by situational elements in the environment in which a person operates. In the same way, Brown [6] examined other “external” factors affecting information behavior, such as organizational structure and the physical environment. Recognizing the existence of other factors, however, does not contradict the psychological approach. Although not explicitly stated, the implicit assumption is that psychological states and processes determine the course of an information-behavior event, but these states and processes can be affected by factors from other dimensions, such as social and environmental ones. The role of the psychological approach is not to investigate what affects cognition, but to focus on how cognition affects information behavior.

In contrast to the psychological approach, the social approach [9] focuses primarily on the study of social, organizational, and political states and processes as an impetus for information behavior. This social approach assumes that the study of information behavior cannot be considered in terms of isolated individuals, or outside a specific context, but should rather focus on the social context and conditions, interaction, and discourse through which human-information interaction occurs. In other words, this social approach views the user as a person who lives and acts within a certain context, rather than a “simple” user of information systems and services such as a consumer. Most studies that adhered to this approach employed theories and frameworks from the social sciences (e.g., diffusion theory, alienation theory, etc.) and studied different communities of end-users (e.g., seniors, experts in a specific domain, engineers, etc.). But, because the social approach assigns prime importance to the socio-cultural context, the results of these studies cannot be generalized across all contexts.

In accordance with Wang, Hawk & Tenopir [33], even if research using the social approach gives a lot of interesting information about the behavior of a community [9, 25, 25, 35], it offers very few descriptive generalizations about information need.

2.2 To Multidimensional Approach

Unlike the other approaches, which focus on one dimension (psychological or social) and thus reduce the complexity, the underlying assumption of our ECS method is that the better this complexity is understood and analyzed, the more relevant the outcomes of the research will be to the design of information systems and services.

Because end-users issued from different cultures do not interact with a system or a product in the same way, cross-cultural differences may cause difficulties in using the product to retrieve information. Taking a holistic approach, Wang, Hawk & Tenopir [33] propose a multidimensional model, which identifies three components: the user, the interface/product/system, and the problem space. Taking a user-centred point of view, the user is the first and the most important element in their model. The problem space is what the user interacts with to obtain wanted information. Between the user and the problem space, there is an interface/product/system, which has been designed to mediate communication between the two. Findings analysis based on this model can provide direct implications for Web design in terms of interfaces and objects and their organization in Web spaces, as well as Web user assistance and training. Other components existing in the Web world such as telecommunication links and programming languages (html, Java, etc.) are omitted from this model because it matters less whether users know about these components in information retrieval or not.

2.3 From the End-Users to Cultural Communities of Users

The user dimension is influenced by dynamic situational factors, such as the particular task, the information need, the culture, and the level of knowledge of the user. In addition, certain individual characteristics influence the human domain, including an individual's cognitive style and affective/emotional state before and throughout the interaction process.

User goals and needs are essential in defining user behavior so that they can be incorporated into a system/product design. User goals and intentions have been theoretically identified as important factors in defining behavior. User intentions are either used as the synonyms of user goals, or as sub-goals that a user has to achieve in the process of accomplishing his or her current search goal. For instance, research in the area of the design of electronic devices related to information retrieval, explores user goals, user information needs, and user information problems [e.g., 2, 15, 16, 13]. These studies suggest that understanding the real goals and needs of the end-users could provide an insight into behavior and strategies and thus system design. Simultaneously, research and studies verify that user goals and intentions should be incorporated into system design to guide effective behavior. For instance, Belkin [2] designed four information system user interface projects in progress. Each project demonstrates a somewhat different approach to interface design, but all have response to user goals, tasks and characteristics in common. They suggested that the design of information systems, including their interfaces, should be based on a multi-level analysis of user goals, tasks and domain views. Later, studies adopted the same point of view in similar ways [22] in order to integrate user goals and needs into system design.

Today, a product and/or a system are created for a large variety end-users who are a not always well identified. So even if the user dimension (from an individual point of view) is very important, the identification of these end-users is often difficult. These days now, a community of users constitutes a “group of users who share the same culture, values, interest, and objectives, and product-related knowledge” [28]. When considering that users share opinions about products and services, and evaluate opinions of peer users, it is important to consider that the organization also reduces the overheads associated with handling users’ enquiries, which result in various organizational benefits, such as time saving and cost avoidance [28].

2.4 The Participatory Point of View to Capture the Influence of Culture

How to obtain a representative evaluation of end-users’ needs issued from different cultures? From a methodological point of view, three types of responses are usually given: (1) some authors seek to transform the verbalizations of users’ needs directly, which implies that the user has the means and skills to formulate such an expression, (2) other authors consider that needs can be deduced for analytical purposes and that by using data interpretation, needs can be inferred, (3) lastly, some authors maintain that need cannot exist as such, outside the bounds of human, historical and societal contexts.

A lot of participatory techniques can be identified: brainstorming, focus groups, participatory design workshops, storytelling, consensus conferences, etc. In all these techniques, the users are invited to speak as freely as possible about software needs, product dissatisfactions, and small technological misfortunes during a session of several hours. The aim is to discover verbally how users exchange information on their interfaces, or to provide interfaces and wait for their collective responses. It should identify the current opinions, preferences or aversions conveyed by these people and could be used to define new specifications. As Salazar-Orvig and Grossen [29] demonstrated, the production of discourse is fairly close to everyday conversation, thus allowing the study of social representations.

However, analysis of verbalizations of these groups requires that, not only the contents of statements by the people are noted, but also relies on the discursive activity of the users to understand how the negotiation of meaning takes place. Indeed, the analysis carried out suggested that people “operate a major activity within a framework to give meaning to the research situation and problems facing them and, according to the framework and interactive work, the subjects' responses vary within a problem, but also to the next problem” [29].

3 Methodological Approach of the ECS

3.1 Main Principles

The ECS method focuses on both the theory of the social construction of technology and on methods to understand users’ requirements and needs [4]. According to the model of Pinch & Bijker [26], the socio-cultural construction of technology is a theory that argues that technology does not determine human action but rather that human action shapes technology. So, the best way to appreciate and design technology is to refer to the socio-cultural context of the technological use. A new technology cannot be designed without understanding how this technology is embedded in its social environment. Comprehension of the social context depends on the methods used to construct and analyze the user's needs.

From a methodological point of view, Brangier [3] and Brangier and Bastien [4] highlighted that user-need cannot exist as such outside humankind, history and the society that generates it. Users' needs are social constructions. Needs are the consequences of complex dealings between users, designers and environments where imitation, learning, co-construction of knowledge and sharing of representations play a crucial function; they involve reciprocal process validations. Need emerges from and through community interactions and through the mediation of language. When designing an artifact, if the user and the designer are not able to resolve their own problems, they will have more opportunity of achieving their objectives by means of collaboration and social interactions. For that reason, needs emerge from collaborative interactions where designers and users reciprocally enrich their understanding by being confronted with the knowledge of others. This knowledge, which finally shapes the representation of needs, can be obtained by using participatory and creative methods. The target of these methods is to explore the creative works constructed by pertinent communities. With participatory methods, verbalizations are produced relating to new forms of cognition that might be useful for users. Briefly, the benefits expected by implementing participatory and creative methods are the development of technologies that are useful and usable for the communities of future users.

3.2 Implications for the Design

Our approach focuses on the social construction of users' needs and it involves the different aspects that designers have to take into account in order to appreciate users' requirements. More precisely, our approach draws on the following ideas:

- An artifact is a social reality before to become a technological reality;
- This reality is not a predetermined construction. It is based on social interactions and is shared (or not) within culture;
- In social contexts, individuals share and disseminate their representations. Each socio-cultural community will develop a flexible interpretation of the technology, its functionality, its aesthetics and its overall use;
- The design of an artifact must be considered from the cultural viewpoint of each relevant community, even if and because, these perspectives are different;
- Because language is the most important way to construct, share and disseminate representations about future technological functions and attributes, analysis of verbalizations produced by different communities must be fundamental.

3.3 Organization of the Expert Community Staff (ECS)

The ECS method takes the form of several groups of experts representing different communities involved in a project and speaking together on the project topic, coordinated by a facilitator [20, 34]. This facilitator uses various media (paperboard, screens, computers, models, storyboards, mock-ups, etc.) to solicit verbalizations and discussions. The expert staff method aims to confront the inter-subjectivity of the group members in order to generate, as much as possible, ideas, needs, goals, and representations of the systems. The overall process of gathering information that lies within the staff method is based on five stages [5]:

- **Stage 1: Defining the communities of practices involved in the project** [11]. Rather than form groups of people representing the general population, the expert community staff method begins with a map of communities potentially targeted by the project. The aim is to gather as much information as possible about people who may be involved by the technology and to determine users' profiles. A community of practice is defined as (a) a community of interest (i.e., a group of individuals sharing a common use which is, more specifically, the users targeted by the product design) and (b) a community of domain knowledge (i.e., users who have knowledge about the uses targeted by the product design).
- **Stage 2: Identifying communities' experts**. The determination of the communities of different practices helps to find the experts of each community. These experts must be recognized as such by members of the community. It means that we need to set up a panel of experts who are representatives of the possible community, who are legitimate and/or recognized. An expert is a valid representative of the community, who could share representations of the future of his community.
- **Stage 3: Organizing and leading each group of staff**. From the panel of experts, each target community is represented by at least one group comprising 4 to 6 people who are filmed for 3 hours to 3 hours 30 minutes. Discussion groups are organized into three phases: (1) Participants are encouraged to speak freely on all the topics related to the project, (2) participants are asked to discuss some media presentations: screens, layouts, or storyboards, (3) individuals are involved in organizing knowledge about the project by doing a kind of card sorting activity together.
- **Stage 4: Analysis of results**. The videos are then analyzed to identify new ideas that will provide recommendations relevant to future product features, systems or services. All the comments made by the experts are transcribed and ideas are extracted from the transcriptions. Obviously, these ideas are numerous, rich and supplied without restriction. The interest of these ideas related to the creation of new features is variable and depends on the project objectives. The content analysis is used to identify the options for the features and is deduced from the uttered needs. These needs are discussed and shared with the stakeholders.
- **Stage 5: Negotiating the consensus**. During a "workshop or consensus meeting", the main objective is to develop consensus with and within the stakeholders, by presenting the viewpoints of each community. The results are discussed and reinterpreted to enhance the objectivity of users' requirements.

4 Discussion

The ECS method involves a variety of pointers to communicate and construct meanings around a project that relates directly to communities of practice. A characteristic of staff is that new ideas are not only born from systems to collect data, but appeared as production of interactions between people, as a co-construction phase involving people, products, as results of communities' exchanges, and as imagined representations of future uses. Thus, the method, using expert staff, really shares the idea that needs are not facts but constructions, not constant or definitive but the result of a long process of collaboration. Need does not exist as such, outside humans, history, or the society that generates it. Need is constructed by social interactions. The ECS method is used to produce data based on the social construction of needs for each community of practice. In other words, our ECS method is a technique of mediation between the socially constructed needs and the stakeholder: the demand, needs, wishes and desires expressed by the staff must be transmitted and the expertise must be mobilized to negotiate with stakeholder.

According to us, the Expert Community Staff can be a new paradigm methodology to help designers to identify the real and effective users' needs, because these needs are socio-cultural constructions that can be extracted from relevant conversations with users. Be that as it may, the discussion on the vast problem of defining "information need" and how "information" can satisfy the end-user is still open [14].

References

1. Allard, S., Levine, K., Tenopir, C.: Design Engineers and Technical Professionals at Work: Observing Information Usage in the Workplace. Journal of the American Society for Information Science & Technology 8, 128–149 (2009)
2. Belkin, N.J.: Anomalous states of knowledge as a basis for information retrieval. Canadian Journal of Information Science 5, 133–143 (1980)
3. Brangier, E.: Besoin et interface. In: Akoka, J., Comyn-Wattiau, I. (eds.) Encyclopédie de l'informatique et des systèmes d'information, pp. 1070–1084. Vuibert, Paris (2007)
4. Brangier, E., Bastien, J.-M.-C.: L'analyse de l'activité est elle suffisante et/ou pertinente pour innover dans le domaine des nouvelles technologies? In: Vallery, G., Amalberti, R. (eds.) L'analyse du travail en perspectives: influences et évolutions. Toulouse: Octarès, Collection "Entreprise, Travail, Emploi, pp. 143–156 (2006)
5. Brangier, E., Dinet, J., Bastien, J.M.C.: La méthode des staffs d'experts de communautés. Orientation théorique, démarche méthodologique et application pratique. Document numérique 12(2), 111–132 (2009)
6. Brown, M.: A general model of information-seeking behavior. In: Griffiths, J.M. (ed.) Proceedings of the American Society for Information Science, 54th Annual Meeting, Washington, DC, Medford, NJ, October 27-3,1 pp. 9–14 (1991)
7. Bruce, H.: A cognitive view of the situational dynamism of user-centered relevance estimation. Journal of the American Society for Information Science 45(3), 142–148 (1994)
8. Byström, K., Järvelin, K.: Task complexity affects information seeking and use. Information Processing & Management 31(2), 191–213 (1995)
9. Chatman, E.A.: Framing social life in theory and research. New Review of Information Behaviour Research, 3–17 (2000)
10. Choo, C.W., Detlor, B., Turnbull, D.: Web Work: Information Seeking and Knowledge Work on the World Wide Web. Kluwer Academic Publishers, Dordrecht (2000)
11. Consalvi, M., Eilrich, L., Sillaume, G.: Rôles et perspectives des bibliothèques numériques dans l'économie de la connaissance: une étude de cas. (online), publishing.nsf/0/FDECF548D12BC30BC12575140048AB73/$file/15h30_CONSALVI_EILRICH_SILL.pdf (2008), http://www.tudor.lu/cms/lisbonne2008/ (accessed March 24, 2010)

12. Dervin, B.: An overview of sense-making research: concepts, methods and results to date. International Communications Association Annual Meeting, Dallas, Texas (1983)
13. Dervin, B.: On studying information seeking methodologically: The implications of connecting metatheory to method. Information processing & Management 35(6), 727–750 (1999)
14. Dinet, J.: Collaborative information search in primary grades: Impacts of domain knowledge. In: Proceedings of IADIS, Cognition and Exploratory Learning in Digital Age (CELDA 2008), Freiburg, Germany (October 13- 15 2008)
15. Dinet, J., Tricot, A.: Recherche d'information dans les documents électroniques. In: Chevalier, A., Tricot, A. (eds.) Ergonomie cognitive des documents électroniques, pp. 35–69. Presses Universitaires de France, Paris (2008)
16. Dinet, J., Favart, M., Passerault, J.-M.: To search information with an Online Public Access Catalogue (OPAC): The impacts of information search expertise on the use of Boolean operators. Journal of Computer Assisted Learning 20, 338–346 (2004)
17. Fidel, R., Davies, R., Douglass, M., Holder, J., Hopkins, C., Kushner, E., Miyagishima, B., Toney, C.: A visit to the information mall: Web searching behavior of high-school students. Journal of the American Society for Information Science 50(1), 24–37 (1999)
18. Ford, N., Miller, D., Moss, N.: Web search strategies and approaches to studying. Journal of the American Society for Information Science and Technology 54(6), 473–489 (2003)
19. Hjørland, B., Albrechtsen, H.: Toward a New Horizon in Information Science: Domain-Analysis. Journal of the American Society for Information Science 46(6), 400–425 (1995)
20. Kitzinger, J.: Le sable dans l'huitre : analyser des discussions de focus group. Bulletin de psychologie 57, 299–307 (2004)
21. Kuhlthau, C.C., Vakkari, P.: Information seeking in context (ISIC). Information processing & Management 35(6), 723–725 (1999)
22. Pejtersen, A.M., Sonnenwald, D.H., Buur, J., Govindarej, T., Vicente, K.J.: Using cognitive engineering theory to support knowledge exploration in design. In: Hubka, V. (ed.) Proceedings of ICED 1995, 10th International Conference on Engineering Design, Prague, August 22-24, pp. 219–229. Czechoslovakia, Zurich, Switzerland: Heurista (1995)
23. Pennanen, M., Vakkari, P.: Students' conceptual structure, search process, and outcome while preparing a research proposal: A longitudinal case stud. Journal of the American Society for Information Science and Technology 54(8), 759–770 (2003)
24. Pettigrew, C.: Waiting for chiropody: Contextual results from an ethnographic study of the information behavior among attendees at community clinics. Information processing & Management 35(6), 801–817 (1999)
25. Pettigrew, C.: Lay information-provision in community settings: How community health nurses disseminate human services information to the elderly. Library Quaterly 70(1), 47–85 (2000)
26. Pinch, T.J., Bijker, W.E.: The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. Social Studies of Science 14, 399–441 (1984)
27. Poikonen, T., Vakkari, P.: Lay persons' and professionals' nutrition-related vocabularies and their matching to a general and a specific thesaurus. Journal of Information Science 35(2), 232–243 (2009)
28. Rowley, J.: Digital consumers: reshaping the information professions. Journal of the American Society for Information Science and Technology 61(3), 125–148 (2010)
29. Salazar-Orvig, A., Grossen, M.: Représentations sociales et analyse de discours produit dans des focus group: un point de vue dialogique. Bulletin de Psychologie 57, 263–272 (2004)

30. Talja, S., Keso, H., Pietilainen, T.: The production of « context » in information seeking research: A metatheoretical view. Information Processing & Management 35(6), 751–763 (1999)
31. Taylor, R.S.: Question-negotiation and information seeking in libraries. College and Research Libraries 29, 178–194 (1968)
32. Yerbury, H., Parker, J.: Novice searchers' use of familiar structures in searching bibliographic retrieval systems. Journal of Information Science 24(4), 207–214 (1998)
33. Wang, P., Hawk, W.B., Tenopir, C.: Users' interaction with World Wide Web resources: an exploratory study using a holistic approach. Information Processing and Management 36, 229–251 (2000)
34. Wibeck, V., Adelsward, V., Linell, P.: Comprendre la complexité: les focus groups comme espace de pensée et d'argumentation à propos des aliments génétiquement modifies. Bulletin de Psychologie 57, 253–261 (2004)
35. Williamson, K.: Discovered by chance: The role of incidental information acquisition in an ecological model of information use. Library & Information Science Research 20(1), 23–40 (1998)

The Necessity of Personal Freedom to Increase HCI Design Quality

Rüdiger Heimgärtner[1], Helmut Windl[2], and Alkesh Solanki[2]

[1] Intercultural User Interface Consulting (IUIC)
Lindenstraße 9, 93152 Undorf, Germany
ruediger.heimgaertner@iuic.de
[2] Continental Automotive GmbH, Siemensstraße 12,
93055 Regensburg, Germany
{Helmut.Windl,Alkesh.Solanki}@continental-corporation.com

Abstract. Creativity needs freedom. This must be considered by project managers and senior managers of HCI design organization. The improvement of quality through job enrichment and at the same time saving costs demands personal skills, expertise, motivation and creativity. Creating task models leads to a qualitative high-grade HCI design. Usage-centered design is based upon a user interface derived from a series of models containing interrelated task cases. Nevertheless, creative processes must have been installed. As creativity needs freedom to increase quality and efficiency in general, this should be widely applied. Job-enrichment achieved through expertise, motivation, creativity as well as thorough better planning is an essential step for creating qualitative HCI design.

Keywords: HCI design, freedom, creativity, quality.

1 Superior HCI Design Requires Creativity

Human Computer Interaction (HCI) - design like any other design activity requires predominantly systematic work. However, there is always a gap in the overall HCI design process, which requires a great deal of creativity in its conceptualization phase and it is almost impossible to bridge this gap with methods or structured approaches [1]. Even if the concept phase needs to be kept as small as possible for the sake of efficiency, HCI design contains process parts that are inherently “creative”. Various approaches and methods try to systematize the design process through organized gathering of information about the use of the system to be designed and structuring the findings to deliver a sound base for the design of the user interface. Presenting the information about usage of a system in form of models provides a very fundamental basis for the creative step towards the design specification.

Usage-centered design is such a model-driven approach for presentation design and interaction design of software [2] and Web-based applications [3]. In contrast to user centered design, usage-centered design is based on formal, structured models that build together a seamless process from analysis phase to the final system. This robust and adaptable process has a proven record of application on a wide range of projects

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 62–68, 2011.
© Springer-Verlag Berlin Heidelberg 2011

from small agile projects [4] to large-scale industrial tools development [5]. For projects coping with complex problems where efficiency, user performance, and safety are crucial it has led to radical improvements and innovations in user task and problem-solving performance [5].

The idea of usage-centered design is to derive a user interface directly from a series of models, and to delay the creative design as far as possible to collect, systematize and structure knowledge about the users, tasks, usage and environment of the system before the abstract models are resolved into a specific user interface presentation. The core of the approach is a detailed task model containing interrelated task cases (use cases in their essential form), which are derived from the roles users play against the system. Based on the task model the canonical prototype, which is an abstract model of the user interface is deduced, which is in the final creative step the groundwork for the detailed user interface [1]. Although the canonical prototype helps to reduce the gap between the abstract model and the real user interface, creativity is still the critical ingredient for an excellent user interface design. This creativity necessary for designing innovative and novel solutions requires training and education, talent and capability plus an environment supporting brainwork and solution finding. This environment is the investment a company has to spend in order to be successful in the marketplace and to acquire a sustainable competitive position by patenting the resulting innovations.

Simplistic or automated attempts to generate a user interface without creative processes have resulted, up to now, in user interfaces to be proved that superior usability can be the result of such an approach.

The innovations in HCI-Design can be roughly divided into the following fields: visual design and aesthetics, interaction, functionality and user interface architecture [1].

The user interface of the mobile phone "Apple iPhone" exhibits a well-known example of an innovative user interface architecture, which not only offers improved functionality, but also radically changes the way one interacts with the device. The key to such innovative user interfaces is creativity [6] of the designer paired with the necessary – but not necessarily project relevant – background knowledge. At the same time the user interface of number one mobile producer Nokia’s Symbian OS is not able to deliver the ultimate user experience what Google’s Android and Apple’s iOS4 are able to and for that reason Nokia has been losing its substantial amount of market share.

2 Creativity Needs Personal Freedom

According to Von Stamm [7], “Creativity is not something where someone who has never worked in that field before suddenly gets this marvelous idea. Creativity is relating a concept to a particular body of knowledge. The existing body of knowledge is as important as the novel idea and creative people spend years and years acquiring and refining their knowledge base – be it music, mathematics, arts, sculpture or design.“ [7]. In this sense, creativity is the process of deriving new ideas from a huge and refined knowledge base. Creativity can be regarded as “the application of knowledge and skills in a new way to achieve a valued goal.” [8].

According to Amabile [9], within every individual, creativity emerges as a function of three parameters: expertise, creative-thinking skills, and motivation. Managers can influence these components positively by specifying workplace practices and conditions.

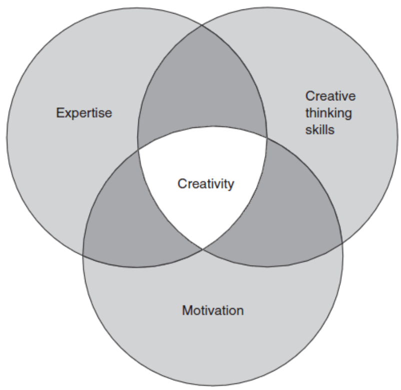


Fig. 1. The components of creativity according to Amabile [9]

Amabile [9] identified six general categories of managerial practice that affect creativity. These are

- challenge (of matching people with the right assignment),
- freedom (of autonomous working which boosts intrinsic motivation and sense of ownership),
- resources (time, money and physical space),
- work-group features (designing mutually supportive teams/groups with a diversity of perspectives and backgrounds),
- supervisory encouragement (praising creative efforts – successful or unsuccessful; encouraging collaboration and communication within the team; creating environment with no worry of failure; creating a culture of “work and play(experiment/explore)” promoting enjoyment), and
- organizational support (mandating information sharing, open communication and collaboration and ensuring that political problems do not foster).

Amabile [10] also looked at whether our feelings affect our creativity. During her research on projects which required creativity and innovation such as new product developments, new processes and solving complex client problems (she characterized all these tasks as ‘knowledge work’) she found that there exists a purely linear relationship: *the greater the positive affect (the happier the person) the greater their creativity* (this reflects the requirement that successful and fruitful thinking needs freedom e.g. based on Aristotle: the thinker needs to be freed from physical work). Other studies from Amabile [10] show that there are positive effects if the person has enough room to do what he or she wants to do.

3 "Personal Freedom"

Both for additional creativity and also for building up required background knowledge, personal freedom outside the everyday project work is absolutely necessary. This available scope / personal space can be used for activities which allow the employee to have more fun during his or her work and help the employee to acquire creative-thinking skills or to improve it. Of course, these activities should also have clearly specified strategic goals that are aligned with organisational goals. In this way, creativity and innovation are supported [11]. The effects of job enrichment on employees are satisfaction, motivation, involvement, and performance, because relieving the employee allows him or her to create more personal freedom through higher performance/productivity [12]; or to use that available scope more effectively than before measured by job diagnostic surveys [13, 14]. This promotes, alongside high internal motivation, high satisfaction and high-quality work performance, the formation of new abilities and skills for the employee, which automatically increases his or her efficiency. This is particularly the case for HCI design based on creativity. Better design and functionality, fitting better the user's needs and tasks, can be achieved through better knowledge about customers and technologies eventually leading to higher quality of HCI design and better usability. Lack of innovation, due to a lack of personal freedom to acquire new knowledge and skills, leads to substantial competitive disadvantages, which is even more true in the context of HCI design. Especially this is the case when the HCI design is gaining more and more importance as a salient feature of product differentiation.

4 Basic Problem: No Freedom Available

As result of overly tight planning which is not foresighted enough, unplanned efforts in development cause resource shortages. Even though under some circumstances, time pressure can enhance creativity, managers do contribute in killing the creativity by planning under the "threshold of sufficiency" or by setting fictitious or unrealistic deadlines to simulate the time pressure eventually. Under such tight circumstances in order to avoid endangering of the project, predominantly the measures to maintain the project on track are carried out and all the work-time is invested in keeping up with the extremely tight time schedules. In this way, the creativity and therefore any scope for innovation is destroyed. In fact the creativity is channelized in other direction e.g. how to get additional resources. Creativity often takes time to explore new concepts, to put together unique solutions, and to wander through the maze of the unknown / unpredictable. In order to achieve innovation, planning personal freedom is necessary as it is anyway impossible to plan 100% resource utilization during the development of the HCI design, as outstanding problems of the concept phase need to be handled and thus unplanned expenditure of time and money occurs.

5 Possibilities for Creating Personal Freedom and Associated Problems

There are various possibilities for creating personal freedom. Building personal freedom through increasing productivity could be accomplished, for example, through the application of methods like job enrichment and/or job engineering [14]. In particular, job enrichment leads to more independence through empowerment and autonomy and in this way to more possibility of choices and freedom, but also to taking more responsibility [15] as well as to higher job performance [16]. Barrick & Mount [16] found that autonomy does play a moderating role on the relationships between personality dimensions (cf. Big Five[1]) and job performance. The emergence of personal freedom within the normal working day is supported above all by healthy workers' high levels of intrinsic motivation [15]. According to [17], work performance is a function of the product of opportunity, capacity, and willingness. For this, securing the health and well-being of the employees is indispensable. This is made possible primarily through a congenial working-atmosphere which supports team work and an acceptable workload through suitable resource planning (resulting in the establishment of a culture of trust [18]). Additionally, resource shortages can be avoided or lessened and personal freedom created via reserving capital to plan extra time. E.g. full time employees are supplied with a basic resource buffer of approximately 20% up to 50%. A proportional creation of personal freedom to support creativity and innovation has already been tried out by many companies and today these companies are the innovation forerunners. Two examples are 3M and Google (cf. HR Policies of 3M [19] & Google [20]). One of the pillars of the "3M Way" was that employees could seek out funding from a number of company sources to get their pet projects off the ground. Official company policy allowed employees to use 15% of their time to pursue independent projects. The company explicitly encouraged risks and tolerated failure. 3M's creative culture foreshadowed the one that is currently celebrated unanimously at Google. Only few companies are as creative as Google. The marvel of Google is its ability to inculcate a sense of creative fearlessness and in the end, the resources and liberty Google entrusts to its employees infuse them with a rare sense of possibilities and obligation. Google's integrated culture comprises four key elements: mission, innovation, fun and reward. At Google there is a 20% rule whereby 20% of employees' time is spent on whatever project they desire [21]. However, the creation of personal freedom always requires some form of investment which initially causes outlays.

---

[1] The "BigFive" consist of (a) Extraversion (e.g., sociable, talkative, and assertive), (b) Agreeableness (e.g., good-natured, cooperative, and trusting), (c) Conscientiousness (e.g., responsible, dependable, persistent, and achievement oriented), (d) Emotional Stability (viewed from the negative pole; tense, insecure, and nervous), and (e) Openness to Experience (e.g., imaginative, artistically sensitive, and intellectual).

6 Solving the Problem of Costs: Prioritizing Measures to Create Personal Freedom

In order to increase its desirability and lower the costs of creating personal freedom, the identified measures need not be introduced all at once, but rather step by step:

Step 1: Planning time for the inventive tasks that need to be carried out within the scope of the project.

Step 2: Creation of additional personal freedom alongside the current project work for innovation which is not connected to the project at hand but can be introduced into future products.

Step 3: Promotion of training and further education e.g. project planning courses for project managers and executives or HCI training for designers and software developers (expertise and knowledge-oriented sustainable personnel management [22]).

Step 4: Job enrichment leads to the implicit creation of personal freedom through autonomy, as the degree of employees' self-effectivity (the satisfaction and feeling of having achieved something and being appreciated) is very high (this means that self-esteem (self-assessment is improved) and finally self-actualization increases [23]).

Step by step execution of the measures saves cost, also because the need for the lower-priority measures is lower when the higher-priority measures have already been introduced. In other words planning time for the inventive tasks in the project supports the building of experience and skills thus of innovation. These lead to a reduction in the measures required for training and further education or job enrichment.

7 Conclusion

Investment in an initial time buffer for creative tasks in HCI design is not only repaid via its effects in the form of higher productivity and product quality with financial gain but also in other ways like exponentially increased loyalty, motivation, or innovation emerging due to the development of synergy effects (e.g. via the feeling of belonging to an organization which is successful for this reason). In this way personal freedom is also created by the employees themselves resulting in creative and innovative HCI design.

References

1. Constantine, L.L., Windl, H., Noble, J., Lockwood, L.A.D.: From Abstraction to Realization: Abstract Prototypes Based on Canonical Components. Working Paper, The Convergence Colloquy (July 2003), `http://www.foruse.com/articles/canonical.pdf`
2. Constantine, L.L., Lockwood, L.A.D.: Software for use: A practical guide to the models and methods of usage-centered design. ACM Press, Addison-Wesley Publishing Co., New York (1999)

3. Constantine, L.L., Lockwood, L.A.D.: Usage-Centered Engineering for Web Applications. IEEE Software 19(2) (March/April 2002)
4. Patton, J.: Extreme Design: Usage-Centered Design in XP and Agile Development. In: Constantine, L.L. (ed.) Proceedings of the First International Conference on Usage-Centered, Task-Centered, and Performance-Centered Design forUSE 2002, Ampersand Press, Rowley (2002)
5. Windl, H.: Designing a Winner: Creating STEP 7 Lite with Usage-Centered Design. In: Constantine, L.L. (ed.) Proceedings of the First International Conference on Usage-Centered, Task-Centered, and Performance-Centered Design for USE 2002. Ampersand Press, Rowley (2002)
6. Schumpeter, J.A.: The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle. Transaction Publishers, N.J (1934)
7. Von Stamm, B.: Innovation, Creativity and Design. John Wiley and Sons, Chichester (2003)
8. Seltzer, K., Bentley, T.: The Creative Age: Knowledge and Skills for the New Economy Buckingham: Demos (1999), http://www.creativenet.org.uk
9. Amabile, T.: How to Kill Creativity. Harvard Business Review, 77–87 (September–October 1998)
10. Amabile, T.M., Barsade, S.G., Mueller, J.S., Staw, B.M.: Affect and Creativity Work. Administrative Science Quarterly 50, 367–403 (2005)
11. Orpen, C.: Effects of Job-Enrichment on Employee Satisfaction, Motivation, Involvement, and Performance – Field Experiment Human Relations, vol. 32(3), pp. 189–217 (1979)
12. Hackman, J.R., Oldham, G.R.: Motivation through the design of work: Test of a theory. Organizational Behavior & Human Performance 15, 250–279 (1977)
13. Hackman, J.R., Oldham, G.R.: Development of the Job Diagnostic Survey. Journal of Applied Psychology 60, 159–170 (1975)
14. Umstot, D., Bell, C.H., Mitchell, T.R.: Effects of job enrichment and task goals on satisfaction and productivity: Implications for job design. Journal of Applied Psychology 61, 379 (1976)
15. Herzberg, F.: The Motivation to Work. John Wiley and Sons, New York (1959)
16. Barrick, M.R., Mount, M.K.: Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance. Journal of Applied Psychology 78, 111–118 (1993)
17. Blumberg, M., Pringle, C.: The Missing Opportunity in Organizational Research: Some Implications for a Theory of Work Performance. The Academy of Management Review 7(4), 560–569 (1982)
18. Miles, R.E., Snow, C.C.: The new network firm: A spherical structure built on a human investment philosophy. Organizational Dynamics 23(4), 5–19 (1995)
19. 3M HR Principles, 3M, http://solutions.3m.com/wps/portal/3M/en_US/us-diversity/diversity/3M/3m-invests-diversity/hr-principles/
20. Google Inc. Ten reasons to work for Google (2005), http://google.com/jobs/reasons.html
21. Byte, X.: Google's answer to investors is 70 20 10 but is not enough, Yeald, http://yeald.com/Yeald/a/35701/google_s_answer_to_investors_is_70_20_10_but_is_not_
22. Pfeffer, J.: Producing sustainable competitive advantage through the effective management of people. Academy of Management Executive, vol. 19, no. 4 (2005)
23. Maslow, A.: A theory of human motivation. Psychological Review 50, 370–396 (1943)

Developing Idea Generation for the Interface Design Process with Mass Collaboration System

Daeeop Kim, Sang-su Lee, Seungwoo Maeng, and Kun-Pyo Lee

Industrial Design, KAIST, Yuseonggu, Daejeon, South Korea
{up4201,mf_largead,maengoon,kplee}@kaist.ac.kr

Abstract. This study is about an Internet-based web service realization to collect data using mass collaboration to put in place Organic User Interface (OUI). As mass collaboration draws attention as a means of collaboration, and its effects are proven, a number of studies have suggested applications and problem-solving skills based on mass collaboration. This study introduces cases of experimenting with mass collaboration, and suggesting an example of system structure design that cares for interaction design.

Keywords: Mass Collaboration, Interaction Design, Organic User Interface.

1 Introduction

1.1 Research Background

The Internet era called web 2.0 has come since the 'dot com bubble' was burst. web 2.0 is a web environment which supports openness, sharing and cooperation technology.[1] Social computing technology representing the web 2.0 era reduces transition costs arising from managing an organization.[2] Social Computing has instantaneity like emails and lays a foundation for community of practice where associated human networks share and reproduce information and initiate the formation of groups, allowing for organized works done as businesses do without the need of traditional groups.[3] Nonaka said that knowledge spiral of knowledge management helps transit of tacit knowledge and explicit knowledge to cause knowledge as a whole to develop via a knowledge spiral in the community. Practical and specialized knowledge like design also can be inherited and evolve by forming a social computing-based community.[4]

Like those of professional designers, works of amateurs participating in Flickr.com and Make.com are purchased by those who are fond of their works.[5] Communities of Practice like LINUX have cases showing that it is possible to develop huge software architecture Operating System (OS), which requires a group as enormous as existing businesses and huge operating costs to maintain such group via an instantaneous communication method and open source.[6, 7] Such participatory sharing is called *Cooperate 3.0*, which is the most advanced form of mass collaboration, classified by knowledge sharing and communication dimension. Cases in this category share and evolve contents sources. [8]

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 69–76, 2011.
© Springer-Verlag Berlin Heidelberg 2011

As an example of success based on slight effort of participants, Wikipedia.org shows that a small unit of knowledge participating in totally open contents source can become a gigantic, live, and effective encyclopedia.[7] Keeping this social phenomenon in mind, Bruns called those who create via collaboration in many production units including software as *'Produseage'* and explained that the community without organization creates special products and innovation beyond personalization. Is it possible to apply *Cooperate 3.0* to the interaction design process as well then? [9]

2 Interaction Design Process and Mass Collaboration

2.1 Traditional Design Process

As design methodology belongs to tacit knowledge, hard to systemize literally, it is difficult to generate a generalized definition unlike other field of study. Nevertheless, design methodology has been defined by many researchers in terms of its types, characteristics, and objectives and has a set of common properties.

First of all, projects have established according to design objectives. Characteristics of design process classified commonly by many researchers including Ulrich are 1) defining design objectives, 2) finding and evaluating solutions, and 3) realizing the selected design plan. These design activities require key information for design process.[10] Such information is essential in determining elements to be coordinated to produce a design and divided into two types: Needs Based Information (NBI) and Solution Based Information (SBI). Generally, the SBI for the selection of final design is composed of specialized information and formed by a designer and a skilled engineer who can support him or her. In the event of any problem in this course, the designer and the engineer try to revise the draft design as possible or otherwise solve the problem.

In the meantime, interaction design products require a process of collecting ideas of users and correcting problems, if any, in the design process. Even though the NBI has been collected via questionnaire survey, some have suggested that limits of questionnaire survey prevent a design success. [11, 12, 13] If types of questionnaire and samples are focused on certain population, depending on the method of survey it is hard to collect and obtain appropriate information on how to use a product and share experience. Accordingly, since the 1990s, user observation has been used widely for this purpose, which makes use of qualitative face-to-face interviews of users by the designer. This method requires much time and fewer participants, raising the issue of objectivity. Nevertheless, qualitative data obtained by this method have made a significant contribution to the innovation of design.[14] As concentrated questionnaire and samples began to be relieved to some degree with the help of online culture lately, solutions via active online participation have been explored, but still there is a growing need for a way to collect in-depth information using social computing.

2.2 Characteristics of Mass Collaboration

How does social computing-based design approach work and basically what kinds of functions does it consist of? Prior to establishing a system, we examined characteristics of shared functions of existing ones. Using Google Scholar, we searched related studies with mass collaboration and crowdsourcing as keywords and collected 16 studies among which six theses were closely related. Focusing on websites introduced in these papers, we found 46 similar websites using Xmarks.com and summarized characteristics of each website in four common keywords as below.

Social Program based Communication: *presentation of new ideas –both NBI required at the beginning of design process and SBI needed at the stage of problem-solving used existing social network as a tool to discuss in detail information necessary in design process or provided similar functions.*

Common Creation Tool: *Open-source based websites provide shared platform which users can download or common creation tool to share production methods. As a result, users can present and combine excellent designs in detailed aspects such as form and function.*

Instruction: *there were web pages which provide examples and methods in regards to idea presentation, operating rules, and production method, as well as detailed description about structures. Also, forums were available via shared bulletin board. Information was open or provided exclusively in this case.*

Evaluation: *presented design ideas were evaluated in terms of oral statement, sentences, and preferred ideas. Scoring votes were possible, and plugins were provided to post evaluation on Facebook or Twitter.*

Taken together, most activities in the design process can happen online. However, it was hard to examine whether or not design activities were performed consecutively or in a linked manner according to a set of goals. This is because participants simply present their ideas or solve minor problems, but they cannot perform activities in a linked way, as they are not trained to solve a problem in a uniform manner as a designer does. With reference to design activities, there are three major modules. This is the case not only in design methodology but also in the innovation process. (See Figure)

Among websites mentioned above, some make use of design process in a broad meaning, and they are divided into two categories; one category consists of open source websites that contribute to the production of contents, and the other category covers websites for the purpose of customizing or modifying a product by reflecting ideas of participants in the substantial design process.

These websites serve as a platform to produce contents via participation or share existing contents and initiate new contents on a separate convenient platform. As one of the *Collaboration 3.0* type websites, Wikipedia.org, for example, allows participants to read contents and produce new contents using Media Wiki, a PHP-based web language platform.[8] Another example includes skin of well-known blogs or open platform-based program Firefox and its Plugin. iPhone App of Apple is a good example, too. These websites are characterized by the utilization of existing design process models to use a practical program available in the computing environment. Although they are applied to the CSCW (Computer Supported Cooperative Work) only, they also can be used to other design process in the future, as predicted by Bruns.[9]

In addition to websites promoting participation of users, we collected theses reporting web-based user participation with the following characteristics:

A. Excellent in collecting needs of users

Generally, theses based on participation of the public have been written with an expectation of collecting needs of users seamlessly. Participation of users was found to be far more effective in collecting needs of users than existing method, as expected.

B. Easy to generate contents from collaboration

We also found cases using collaboration such as crowdsourcing-based translation or word-processing. Clearly, these cases indicate that user participation is helpful quantitatively rather than qualitatively.

C. As specialized as professionals

In regards to design ideas, public participation ratio is high. Some experts point out that there are many different views and innovative ideas presented which they did not predict.

D. Difficult to guarantee the quality of users' opinions

Despite advantages mentioned above, problems were raised: most of the opinions presented by users were similar or in-depth results were not generated. As more people participated, more cases were found useful for design, but with lower frequency.

E. Success dependent on participants

Attitudes of participants were considered problematic. Frequent problems include unfaithful attitude or inappropriate logic, which led to unnecessary works in design process.

Literature research showed that most theses viewed collecting ideas and discussion online as highly positive. In particular, the NBI necessary at the initial stage of design consisted of various ideas from many users, and a number of studies have reported that their participation was very active. However, the SBI to realize ideas in design process was reported as extremely low. More important, this type of information is required by designers and engineers in the design process and a core technology for survival features of designers in that this piece of information is provided by trained designers only. [15]

Interestingly, Hagen claimed that NBI could be solved by social computing in the development of software, and SBI by open-source method. He claimed that software development is underway in many successful web-based services using "Iterate It" and "Emerge It" for SBI and "Source It" and "Open It" for NBI. Some suggest that even realization most difficult in product design is also possible through collaboration when circumstance permits.[16]

3 Implementation

3.1 Design Object

Traditionally, design activities have been defined as behaviors of professional designers to develop a set of processes to produce a product or service in line with the goals of stakeholders and thereby make a design. In addition, design activities can be summarized into Goal Definition, Solution Suggestion, Prototyping (Modeling) and Evaluation, and Implementation.

For Organic User Interface (OUI), its concept has been defined, but still it is necessary to collect activities of users and analyze most 'desired gestures' for concrete realization, because it is useful to use natural bare hand activities in order to develop interface gestures fit to 'mental model of user' most important for interaction design. Comprehensive user participating survey and design proposal are necessary to examine what activities can be used for the OUI and what is the most applicable deformation-based interface to prevent problems arising when engineers and interface realize an interface from their own perspectives. In this respect, the author believe that this topic can help users directly propose a prototype of an interface via collaboration and, in turn, researchers realize it, which is likely to work for better OUI construction. As a result, a system can be installed according to characteristics of activities by design activities and processes.

The system realizes the following:

Suggestions on bare hand gestures fit to WIMP-based interaction previously presented by users are received. (texts, questionnaires, video clips, photos, etc.)

A list of collected bare hand gestures is presented for preference evaluation.

Examples of flexible media are reviewed, and suggestions on 'products to apply' are received.

Ideas on whole new type of OUI-based products are presented.

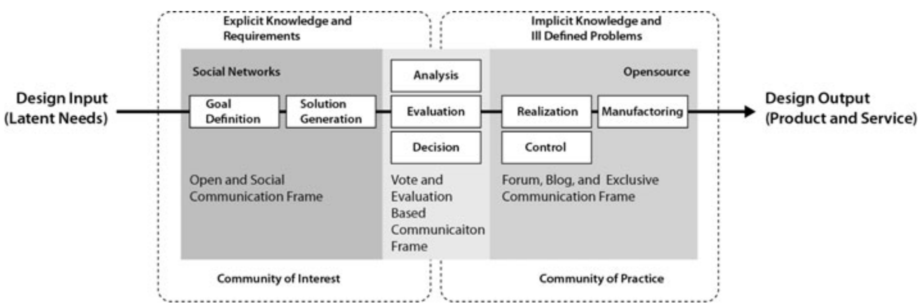


Fig. 1. Figure of the Design System

3.2 Expected Problems of System and Alternatives

Many studies reported cases where general users participated in design process and presented better problem perception and solutions than designers do. Nevertheless, such design ideas were found to have problems at the stage of implementation. While professional designers can suggest the direction for production with specialized views in the process where ideas are concretized through various experience, most ideas of users were fragmentary and minor, inclined to solve problems from their own views. In other words, participants were found to be helpful for the NBI, but not so for the

SBI necessary for actual implementation. To address this problem, we propose paper as a tool of basic prototyping for implementation. Paper is easy to find, has been used in previous studies on key gestures of OUI, and is suitable to suggest gestures without actual implementation.

Things were similar in other websites where collaboration was used to solve problems. Only four communities discuss the SBI among 46 websites found by the author. The fact that hobby-like works are shared only in these communities shows that participating in the SBI is relatively difficult. Therefore, requiring proper level of technology or making difficult technology easy via system is expected to promote mass collaboration for design. In this study, we expect that this problem will be solved by using paper prototyping.

Another problem was that perception and participation in design issues varied qualitatively among participants. Shirky explained this problem as power function.[3] He pointed out that 75% of open source software uploaded at Sourceforge.org is not downloaded. This means that although those who undergo professional design training also take part actively in design process via mass collaboration, the number and the quality are unlikely to be uniform. Surowieki also emphasized that voluntary participation and mature judgment as well as freedom of participants are required to allow participants to determine successful collective intelligence at 'Wisdom of Crowds.' However, many regulatory problems including vandalism expected in design process have been neither found nor discussed academically yet.[17]

Any model to solve this problem should have a convenient interface that allows easy communication so attract a number of users into the system. Shirky suggests three keywords for practice for the revival of mass collaboration at the beginning. [3, 17] This means incessant communication with participants and active supports from the designer rather than experiment. Therefore, instead of making a new system for experiment, existing system which is linked to communities with many users is more appropriate for this study. In this study, we used a system provided by Ning.com that supports communities professionally. Although it is not free, Ning.com provides individual blogs easier to customize and flexible to link to community channels, Facebook, and Twitter, compared to free open tools. It is also easy to construct polls and surveys for design evaluation.

Thirdly, we should think about how to reward participants properly. A series of design process have various instruments for financial profits. For example, Intellectual Property instruments or patents are a right claim for bother designers and businesses to get legal protection. However, the only profit from sharing of design contents including software is emotional satisfaction. Baldwin et al. explained profits of these open innovation groups by emotional satisfaction from contribution, the desire to learn more, and upgrade of status within the community.[18] Therefore, certain design topics maximize emotional satisfaction of participation beyond such non-capital situations, and there is a need for a study on artificial devices for this purpose. In this study, we made efforts to help participants think that they were making a contribution to technical development for the future. No sample was developed, but participation of users would serve as an idea on samples. Feedbacks were continuously given about the prototype produced as a result of implementation of the study.

4 Conclusion

Mass collaboration-based design approach is still at its embryonic stage. Community of practice consisting of non-expert members in specialized areas suffers from many problems such as incomplete opinions or prototype, but helps find specialized information necessary to solve problems with design realization. Currently, most of the social network services used by many people have developed successful interface via lots of failures and repetition by means of quick open-beta, reducing errors over time. To prevent failures of websites, error-correcting costs must be paid, and participation of various people is used to save such costs.

This study examined whether recent trends have an effect on design process by collecting and investigating websites providing mass collaboration-related services and related literature and then developed a common model to install a website fit to an OUI project. It will take time to determine if this study is successful or not, but in case that a large-scale study is necessary to develop a user-friendly interface, approach like this is more likely to be effective and efficient than existing ones.

Acknowledgement. This research was supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R33-2008-000-10033-0).

References

1. O'Reilly, T.: What is Web 2.0 (2005), `http://oreilly.com/web2/archive/what-is-web-20.html` (cited September 25, 2010)
2. Bencler, Y.: The wealth of networks: How social production transforms markets and freedom. Yale University Press (2006)
3. Shirky, C.: Here comes Everybody: The Power of Organizing Without Organizations. Penguin Press (2008)
4. Erden, Z., von Krogh, G., Nonaka, I.: The quality of group tacit knowledge. The Journal of Strategic Information Systems 17(1), 4–18 (2008)
5. Tapscott, D., Williams, A.D.: Wikinomics: how mass collaboration changes everything, p. 324. Portfolio, New York (2006)
6. Lakhani, K., Von Hippel, E.: How open source software works. Research Policy 32(6), 923–943 (2003)
7. Mockus, A., Fielding, R., Herbsleb, J.: Two case studies of open source software development: Apache and Mozilla. ACM Transactions on Software Engineering and Methodology (TOSEM) 11(3), 309–346 (2002)
8. Dutton, W.H.: The Wisdom of Collaborative Network Organizations: Capturing the Value of Network Individuals, University of Oxford (March 26, 2008)
9. Axel, B.: Blogs, Wikipedia, Second Life, and Beyond Peter Lang (2008)
10. Ulrich, K.T., Eppinger, S.D.: Product design and development, 4th edn., 368 p. McGraw-Hill Higher Education. xv, Boston (2008)
11. Cooper, R.G.: Winning at new products: accelerating the process from idea to launch. 3rd ed, vol. xiii, 425 p. Perseus, Reading, Mass. Great Britain (2001)

12. Franke, N., Shah, S.: How communities support innovative activities: an exploration of assistance and sharing among end-users. Research Policy 32(1), 157–178 (2003)
13. Griffin, A., Hauser, J.R.: Sloan School of Management., The voice of the customer. The International Center for Research on the Management of Technology, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, 38 p (1991)
14. Sanders, E.B.N.: Product Development Research for The 1990's. Human Factors and Ergonomics Society Annual Meeting Proceedings 37, 422–426 (1993)
15. Poetz, M., et al.: The value of crowdsourcing: Can users really compete with professionals in generating new product ideas? (2010)
16. Hagen, P., Robertson, T.: Dissolving boundaries: social technologies and participation in design. In: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open, vol. 24(7), pp. 129–136 (2009)
17. Surowiecki, J.: The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations, 1st edn., vol. xxi, p. 296. Doubleday, New York (2004)
18. Baldwin, C., Von Hippel, E., School, H.B.: Modeling a paradigm shift: From producer innovation to user and open collaborative innovation (2009)

The Essence of Enjoyable Experiences: The Human Needs

A Psychological Needs-Driven Experience Design Approach

Joonhwan Kim[1], Sung Park[1], Marc Hassenzahl[2], and Kai Eckoldt[2]

[1] Samsung Electronics, 416 Maetan3, Yeongtong, Suwon, Gyeonggi 443-742, Republic of Korea
[2] Folkwang University of Arts, Universitätsstrasse 12, 45141 Essen, Germany
{joonhwan.kim,sung76.park}@samsung.com, {marc.hassenzahl,kai.eckoldt}@folkwang-uni.de

Abstract. A huge shift in design in the industry has widened the design scope from pursuing usability and visual attraction to covering user's comprehensive experience. One of the most important aspects of the user experience design is providing positive and enjoyable experience to the users. While both tangible and intangible approaches are important, only a few practical studies have focused on the intangible aspects such as emotion and human needs. This paper describes the importance of the fulfillment of the user's needs for differentiated enjoyment of user experience design, and suggests a practical design method. The authors propose an experience design process and method, which helps to generate innovative design concepts based on the user's psychological needs.

Keywords: Psychological human needs, Enjoyable experience, Positive experience, Enjoyment, User experience, Design method, Service design, User-centered design.

1 Introduction

Trends in human-computer interaction have been shifting toward consideration of user's experience. A huge shift in design in both the industry and the academia has widened the design scope from pursuing usability and visual attraction to covering user's comprehensive experience. One of the most important aspects of the user experience design is providing satisfaction and enjoyment while the users are interacting with the product, system or service; in short, providing positive and enjoyable experience to the users.

While both tangible and intangible approaches are important [1], only a few practical studies have focused on the intangible aspects such as emotion and human needs. The problem is that the intangible aspect in enjoyable experience is a very broad concept; models often describe experience as sensual, emotional, motivational and temporal [2], or physio (e.g., touch, taste) and socio (e.g., relationship with others) [3]. Yet such models are often not comprehensive and not supported by empirical data. Gaver and Martin compiled a list of needs, such as novelty, surprise,

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 77–83, 2011.
© Springer-Verlag Berlin Heidelberg 2011

diversion, to influence the environment, to extend knowledge and control, intimacy, to understand and change one's self, and mystery [4]. Hassenzahl suggested manipulation, stimulation, identification, and evocation as important needs in the context of interactive products [5]. However, the needs mentioned in those studies are often not comprehensive or not empirically examined.

A unifying needs model by Sheldon et al. [6] integrated many other prominent theories and identified ten salient needs based on empirical evidence: autonomy, competence, relatedness, stimulation, influence, security, physical thriving, meaning, self-respect and luxury (see Table 1). They asked participants question about what is satisfying about satisfying events, and showed that the degree of need-fulfillment was positively related to the intensity of the positive effect. They also concluded that the ten needs are relatively independent from each other; that is, it is possible to discriminate between experiences.

Table 1. A list of intangible aspects based on a review of prominent need theories

Sheldon, et al. (2000)	Gaver & Martin(2000)	Jordan (2000)	Hassenzahl (2003)
Autonomy – Independence			
Competence – effectance	To extend knowledge and control; to influence the environment	Psycho	Manipulation
Relatedness – belongingness	Intimacy	Socio	
Influence – popularity		Socio	Identification
Pleasure – stimulation	Novelty, surprise, diversion, mystery	Psycho	Stimulation
Security – control			
Physical thriving –bodily		Physio	
Self-actualizing – meaning	To understand and change one's self,	Ideo	Evocation
Self-esteem – self-respect			
Money – luxury			

Being inspired by user’s fundamental psychological needs is an interesting approach in the experience design process, particularly for developing innovative new concepts.

This paper describes the importance of the fulfillment of the user’s needs for differentiated enjoyment of user experience design, and suggests a practical design method. The authors propose an experience design process and framework, which helps to generate innovative design concepts based on the user's psychological needs.

2 Methods

2.1 Redefining Human Needs

Based on the ten psychological human needs outlined by Sheldon et al., some modifications were made to the more practically applicable method. The first five in Table 1 are the ones already addressed by other user experience models. We further added security, due to its link to usability. Luxury was excluded, due to its marginal role in [6]. Self-respect, although a distinct need in [6], was excluded, because it could be understood rather as an outcome of need fulfillment than a need in itself. Additional needs were added to patch the potential "holes". The most obvious was collecting and preserving meaningful things, a need mentioned by Reiss [7], but not covered by Sheldon. Some modifications were made to be the more practically applicable method (e.g., Competition). It is important to note that our set of needs is not meant as a definite selection; rather we aimed at a concise, streamlined list, covering most of experiences without being overly complex.

In conclusion, six main needs were selected: autonomy, competence, relatedness, stimulation, influence and security.

- Autonomy is feeling like human is the cause of her own actions rather than feeling that external forces or pressures are the cause of her action: *"I can do what I want, the way I want it"*.
- Competence is feeling that human is capable and effective in her actions rather than feeling incompetent or ineffective: *"I am good in what I do"*.
- Relatedness is feeling that human regular intimate contact with people who care about her rather than feeling lonely and uncared of: *"I feel close to the people I care about."*
- Stimulation is feeling that human get enjoyment and pleasure rather than feeling bored and under-stimulated by life: *"I am experiencing new activities.*
- Popularity is feeling that human is liked, respected, and has influence over others rather than feeling like a person whose advice and opinion nobody is interested in: *"I have impact on others"*.
- Security is feeling safe and in control of human's life rather than feeling uncertain and threatened by her circumstance: *"I am safe from threats and uncertainties"*.

Additionally, four sub needs were defined: physical thriving, meaning, competition and collecting.

- Physical thriving: *"I am healthy and physically active"*.
- Meaning: *"My life, my activities have a deeper meaning"*.
- Competition: *"I am better than others"*.
- Collecting: *"I am collecting and preserving meaningful objects"*.

2.2 Suggesting Experience Patterns

Experience design requires bridging the gap between abstract needs and concrete product, service and activities. To enhance the creative concept generation, 17 experience patterns are suggested for the six main human needs: *Keep a secret* for Autonomy; *Homemade* and *Improving, seeding & levering* for Competence; *Shared consumption, Together alone, Being a part of it, Feeling close, Mind reading, Mysterious coincidence* and *Share a secret* for Relatedness; *Serendipity* and *Voyeur* for Stimulation; *Being a hero* and *Image* for Popularity; *The menu, Alpha wave* and *Base camp* for Security.

Experience patterns are always related to particular needs, that is, a pattern is a generic way to fulfill a need. The experience patterns bridge the gap between a need and an activity or product. They condense positive everyday experiences to minimal set of crucial elements. This set is sufficient to explain why people enjoy these experiences and highlight the essence of a whole class of experiences.

Each experience pattern consists of:

- Activities with their sequences if necessary,
- Related thoughts and feelings,
- General rules, which bound and shape the experience,
- Important potential problems and their sequences.

An example of the experience patterns is written below.

• *Name: Keep a secret*

• *Need: Autonomy*

• *Description and insights:*

Many people have "true" secrets. Revealing them would have severe consequence. It would disappoint people they love, they might lose their job or face other severe threats to existence. In this case keeping a secret is a necessity. However, there are also situations where keeping a secret is of its own value.

Having secrets supports autonomy. The self, individuality, is not a closed, well-defined entity. It consists of self-knowledge, but also of other people, possessions and so on. Especially people in close relationships experience a blur of the boundary between one's self and the close partner. To maintain information related to oneself ("self-knowledge", e.g., things done, experiences made, people one knows), which is only known to oneself, creates a boundary between the person and others. By this, it reminds people of their individuality and helps them to define themselves. This is supported by the observation that secrets are an important part of children's development. Although "keeping a secret" is positive, it may create mixed feelings because of the tension between the pleasure of having a secret and the guilt of keeping it from close others.

Self-defining secrets only work if they remain a secret. Even the fact that one has a secret is not to be revealed. Any plain signifiers of a secret (e.g., locked drawers or boxes, password protected parts of computers) must be avoided.

Self-defining secrets are rarely "true" secrets. If they become revealed accidentally, others are often unable to understand why this information was kept as a secret at all. Given the role of the secret, this is easy to be understood: it is keeping the secret per se and not the content of the secret that counts.

• *Stories:*

John is a busy person. He has a demanding job, hundreds of other things to attend and an active social live. Occasionally, he feels as if he as a person dissolves. He then sneaks out in the afternoon, goes to a café and just reads a newspaper. He never tells anybody and he deliberately goes to a part of town, where he is certain not to meet anybody. The free hour remains his little secret.

Bernhard has this hard to explain weakness for watching gory horror movies. His wife, Anna, detests this and neither understands nor easily accepts Bernhard's interest in those movies. To avoid constant conflict, Bernhard agreed to fight this "bad habit." He officially threw away his complete horror DVD collection. But, once in a while, when Anna is out, he takes his favorite DVD – the one he managed to save from the garbage – from a secret place and watches it. When Anna comes home and asks Bernhard about his evening, he says "Just wonderful!" and grins. He feels at the same time more alive and a little guilty.

Because of their higher specificity, patterns might not match each and every domain/product. In some cases, the domain/product itself already points at a particular pattern, which in a way "lends itself to application" because of its similarities to the domain at hand. The results may not always be viable, but there is a fair chance of developing interesting novel ideas, which may further lead to innovation. The experience patterns helps the designers match observed behavior and pain points with the proper human needs correctly and suggests the design directions with more actionable details.

2.3 Building a Design Concept Generative Method

Based on the human needs and their matching experience patterns a design concept generative method is built, which consists of five steps.

Step 1. Determine an activity or a product to enhance with experience. Collect all available information about users and their behaviors.

Step 2. Choose a need. Typically, an activity or product suggests a need. (For example, families spend their quality time in front of the TV. Are there any potential ways to enhance the feeling of relatedness while watching TV?)

Step 3. Identify a need-related, applicable experience pattern to shape the experience. Look out for similarities between the activity at hand and the pattern to determine applicability.

Step 4. Contrast a typical current experience with the experience suggested by the chosen experience patterns (Status-Quo-Analysis)

Step 5. Improve the current experience through the activity or product determined in Step 1. Determine ways to shape the experience as suggested by the pattern through a product's functionality, content, presentation, and interaction with users and other products or services.

In Step 2, forcing a need upon an activity or product in is also possible. It may lead to a new insight. (For example, Routines and relaxation are pleasurable as well as exciting and stimulating in TV watching experience. Are there any potential ways to enhance the feeling of security while watching TV?)

3 Results

Using the suggested design method, the authors made various practices and design concepts. In this paper, a couple of concepts are introduced.

(1) The secret digital photo frame. The digital photo frame tends to be located in public places (e.g., office). This may be annoying for users because the photo frame displays personal memories in a public place without filtering. Among the ten human needs, *autonomy* and *memory* are selected to solve this problem. With the "*keeping the secret*" experience pattern, an unique control method is suggested for the private photo viewing, which is turning the photo frame upside down, when no one is around in the office.

(2) The mobile phone as a smart secretary. When the user is not available to answer the phone, the benefit of the mobile phone significantly reduces. Among the ten human needs, *relatedness* is selected along with the "*mind reading*" and "*feeling close*" experience patterns to solve this disconnection between the people. The improved experience includes the caller recognition functionality of the close people and automatic reply when the user is not available to answer, thus synchronizing the user's schedule and context-aware technology.

4 Discussion

This suggested framework and method enables the designers to generate experience designs from both concrete theoretical background and designer's creativity. The human needs and the experience patterns mentioned in this paper capture important aspects of experience. They can spark and guide designer's creativity and the way of thinking.

There is room to define more psychological human needs and to develop the experience patterns further. The suggested experience patterns also can be evaluated and further defined.

In an industry and academia where the authors belong, the suggested design approach is used as a tool to generate the actionable design directions, particularly for the complex and comprehensive experience design such as device to service cross-domain design. There is a fair chance of developing interesting novel ideas, which further leads to innovation.

Acknowledgements. This is based on the outcome of the collaborating project between Samsung Electronics and Folkwang University. Authors would like to thank Sungmin Yoo, Younhee Rho, Minjung Park and Stephanie Heidecker for their support and excellent work.

References

1. Blythe, M.A., Hassenzahl, M., Wright, P.C.: More Funology. Interactions 11(5), 36–77 (2004)
2. McCarthy, J., Wright, P.C.: Technology as Experience. MIT Press, Cambridge (2004)
3. Jordan, P.: Designing pleasurable products. An introduction to the new human factors. Taylor & Francis, London (2000)
4. Gaver, W.W., Martin, H.: Alternatives. Exploring Information Appliances through Conceptual Design Proposals. In: Turner, T., Szwillus, G. (eds.) Proceedings of the CHI 2000 Conference on Human Factors in Computing , pp. 209–216 (2000)
5. Hassenzahl, M.: The thing and I: understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (eds.) Funology: From Usability to Enjoyment, pp. 31–42. Springer, Heidelberg (2003)
6. Sheldon, K.M., Elliot, A.J., Kim, Y.: What is Satisfying About Satisfying Events? Testing 10 Candidate Psychological Needs. Journal of Personality and Social Psychology 80(2), 325–339 (2001)
7. Reiss, S., Havercamp, S.M.: Toward a comprehensive assessment of fundamental motivation: Factor structure of the Reiss Profiles. Psychological Assessment, 10, 97–106 (1998)

A Culturally Driven Approach for the Development of Innovative User Interface Design Concepts

André Liem

Department of Product Design, Norwegian University of Science and Technology, Kolbjørn hejes vei 2B, 7491 Trondheim, Norway
andre.liem@ntnu.no

Abstract. This article attempts to argue that social, political and economic perspectives supported by a cultural understanding of societies and regions are the cornerstones for a more comprehensive external analysis. Extreme trends and developments in nations´ political, economical and social situation are a source for innovation in the development of user interface designs. As social, political and economical developments in a society are difficult to change overnight; various case studies have illustrated the potential role of interface design in improving the negative aspects of these developments, usually represented by extreme cultural trends. A bottom-up analysis of the case studies resulted in the six (6) preliminary categories, which function as a guide for a broader approach in terms of future external analysis and goal finding. However at this moment, it cannot be guaranteed that design concepts generated from a cultural difference perspective are more innovative than others.

Keywords: Cultural Dimensions, External Analysis, Radical Innovation, User Interface Design.

1 Introduction

From a product sustainability perspective, incremental improvements will not suffice anymore. Radical or systemic innovation is needed, whereby a change in the approach in the searching for new solutions is essential [1].

In a world where interface design is both more in demand and profitable than ever before Culture-based user interface design is a new field in interface design, crucial for designers of all disciplines [2]. Reiterating the importance of “Technology Push” and “Need Pull” in search of innovation, many case studies discuss the effects of cultural differences on artefacts or even design of some culture-specified products, however, it is not easy to find models or processes about connecting culture to designing for radical innovation. The closest research activities in this area pertains the works of Marcus [3], where user interface components have been mapped against Hofstede’s cultural dimensions. This mapping then has been translated to a number of patterns for each cultural dimension. This is followed up by the development of culture-oriented human machine systems [4], whereby a detailed model has been applied to analyse intercultural variables and cultural factors (dimensions) in a systematic process. Findings of this process, which can be different components of the design, are then integrated with design requirements (Figure 1).

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 84–93, 2011.
© Springer-Verlag Berlin Heidelberg 2011

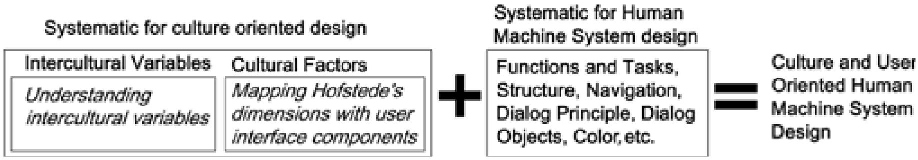


Fig. 1. Integrative Approach of the Culture-Oriented Design [4]

When widening the scope to designing products in general, it is concluded that one of the key attributes that distinguishes breakthrough products from their closest followers is the significant value they provide for users [5]. After all, as Drucker has pointed out, "customers pay only for what is of use to them and gives them value"[6].

Sanders and Simons [5] identified 3 types of values related to co-creation, which are inextricably linked. These values are monetary, use /experience and societal. Bringing in the cultural component, relationships between culture and design have been viewed from different perspectives in literature. From a value creation perspective, three main aspects have been suggested, which make culture an important parameter in product design; “Usability and Ergonomics”, “Business Advantages”, “ Social Sustainability and moral values” [8]. However, referring to Cagan and Vogel’s positioning map [5], culture has not been addressed as a value creation dimension.

2 Methods and Paradigms for Innovation

Within the context of value creation and integrated product development, the level of innovative success in formulating an effective product strategy and a design goal is highly dependent on how thorough ”Product Planning and Goal Finding” processes were carried out in the Front-end of Innovation (FEI) [9]. Effective management of unknown and uncontrollable factors in the front end may result in a sustainable competitive advantage. In this sense, the focus on the front end is mainly one of opportunity identification and analysis [10]. Hereby, both internal and external sources are important for idea development and goal finding, but, the designer’s approach towards the execution of the external analysis determines the level of innovation targeted [11].

In practice, external analysis focuses mainly on market, competitor’s and stakeholder’s analysis, leading to incremental innovation, where new products were created for existing markets or new markets for existing products confining itself to current product or service portfolios of the respective company [12]. To achieve diversification / radical innovation, a broader approach towards innovation processes is needed to obtain a maximum number of innovative product and process ideas. However, there has been little research done on the issue so far [13].

This article attempts to argue that social, political and economic perspectives supported by a cultural understanding of societies and regions are the cornerstones for a more comprehensive external analysis, increasing the chances for diversification.

3 Culture and Its Cultural Dimensions

Culture has been defined in a number of different ways because of its multi-dimensional characteristics. For example, culture is cross disciplinary defined as "transmitted and created content and patterns of value, ideas, and other symbolic-meaningful systems as factors in the shaping of human behaviour and the artefacts produced through behaviour." [14]. Hofstede defines culture as the collective programming of the mind that distinguishes the members of one group or category of people from another, based upon a constellation of psychological traits, attributes, and characteristics [15].

Identifying cultural characteristics is difficult because it lacks a robust measure that can identify the implicit levels of culture [16]. In an effort to address this issue, many researchers have dissected culture as a set of 'dimensions' that provide a framework for cross-cultural comparisons of user behaviour. A Model is developed based on a survey of IBM employees in 40 different countries, describing national cultures that entailed four dimensions: uncertainty avoidance, individualism vs. collectivism, masculinity vs. femininity, and power distance [17]. Hofstede and Bond [18] subsequently added a fifth dimension to their model, long-, vs. short-term orientation.

4 A Cultural Perspective on Innovation and Interface Design

Considering cultural driven innovation, strategists and designers should acknowledge that numerous societies believed that their habits, ideas and customs determined the shape of their political and economic arrangements, and were the source of their uniqueness. When Hofstede´s five dimensions [15] were associated with manifestations of cultural difference, thus linking cultural parameters to cultural behaviour, it was observed that the potential for radical innovation is neither biggest in leading nor developing countries. The feeding ground for radical innovation is the understanding of the status quo of a nation's cultural, political, economic and social atmosphere, followed by the acknowledgement that drastic improvements in quality of life, service quality or minimizing inequalities within societies, are almost impossible to be achieved through political governance.

In terms of thinking, acting and communicating, people adopt patterns from living in a specific social environment, normally typified by national culture [19]. As such, culture partially predetermines a person's communication needs, preferences and behaviours. Hereby, the overall patterns and values of a culture are reflected by the communication style of how a person sends and interprets messages.

Mainstream user interface design research focuses on the means by which the user and ICT products interact. It has been discussed whether or not that the interface should facilitate users to use their particular communication styles [19]. However, it has become clear that user reactions become more predictable and understandable when the user's cultural perspective is taken into account [21, 22]. For example, websites need to be designed to accommodate the cultural preferences and biases to increase the culture-based interface in the product [20].

However, a user-centred and culture driven innovation approach with the aim to achieve diversification is difficult to find. This calls for a design-driven innovation approach, whereby service-oriented, strategic design concepts are proposed to solve cultural extremes at the bi-polar scale of Hofstede´s dimensions [19]. Design-driven innovation embraces a *reconstructionist* or *social-constructionist* view of the market [23, 24], where the market is not "given" a priori, but is the result of an interaction between consumers and firms. Hereby, users, designers and other stakeholders need to find new connections to their socio-cultural context by exploring new values and patterns of interaction with the product.

5 Research Focus and Method

Prior to the research questions below, an assumption is made that extreme trends and developments based on Hofstede´s five cultural dimensions, represented by certain nations are unfavourable. A hypothesis is formulated based on the assumption that unfavourable conditions are a potential source for design innovation, as it is difficult and slow to change a nation's cultural, political, social and economical disposition. Related research questions are:

- Can areas for "quality of life" improvement for certain groups in respective societies be identified and elevated through design by mapping case studies from a social, political and economic perspective on a bi-polar scale, supported by Hofstede's "Cultural Dimensions?
- Is there a potential to develop a methodology for strategic goal finding based on social, cultural and political differences on a bi-polar scale?
- Will the development of innovative design concepts then be more radical by understanding the extremities on the bi-polar scale?

A "Multiple Case Studies" research approach was used to gather findings [25]. Sources of evidence were mainly based on literature studies and observations followed by interviews. The analysis of case study evidence was carried out through a procedure of "Explanation Building". Each case study was summarised and tabulated, according to the following topics: "Context", "Cultural Explanation", "Dominant Cultural Dimension", "Design Problem" and "Design Concept". Thereafter, a comparative analysis was done among the case studies.

6 Summary and Analysis of Case Study Results

Seven (7) case studies were summarised and analysed. On the basis of "Context", all case studies illustrate a clear contradiction based on cultural differences. The "Cultural Explanation" shares more in-depth knowledge on how these cultural differences affects society and human interaction from an economic and political perspective. All Hofstede's cultural dimensions were covered by the case studies; however "Power Distance", "Masculine versus Feminine" and "Collective versus Individual pre-dominantly appear and can be recognised as the main building blocks in determining whether a society is capitalistic or social-democratic. The building blocks were also regrouped as “Large Power Distance – Masculine – Individual” and “Small Power Distance – feminine – Collective” and positioned on the extremes of a bi-polar scale. All “Problems” describe a need for improvement in the area of Interaction / Interface Design from one nation´s contextual perspective in comparison to the opposing nation´s political, social or economical state on bi-polar axes.

The following examples illustrate how comparative cultural case studies have led to potential innovative design concepts.

Health Care, Denmark versus USA

Context: Accessibility of health record systems is greater in social democratic societies, such as Denmark compared to USA. In USA, healthcare is not subsidised and an emphasis is placed on the patient’s privacy.

Cultural Explanation: In the Danish society, the cultural ideology is that all should have equal rights to healthcare, especially the weaker in society. However, protection of personal medical records is lesser emphasised in Denmark than the USA, as the former is capitalising on the prevalence of “human trust”.

Dominant Cultural Dimension: Masculine versus Feminine, Power Distance.

Design Problem: How can the American health service be made more efficient and user-friendly in terms of administration from both governing and professional bodies as well as the patient him / herself? How can personal medical records be managed and controlled as to prevent frequent misuse?

Design Concept: A secure and personal website where people can administrate medical records, compare payment models, insurances, make appointments, as well as extract objective medical information. This website is linked to security number and only made accessible to authorised parties.

Education - India versus Norway

Context: The current literacy rate in India is above 60% and growing steadily. In the year 2005, the total public expenditure on education as a percentage of Gross Domestic Product (GDP) for Norway and India were 7.2 and 3.2 respectively. The difference in percentage of tertiary education is small compared to the literacy rate between the two countries. India is the 3rd largest TV market in the world and the cost of single television set is a fraction of yearly salary of semi-trained teacher.

Cultural Explanation: India is facing a tough challenge to increase the literacy rates and improve the quality and accessibility of their education system. The education system is hierarchical and competitive at all levels. An emphasis is placed on performance and the better students are favoured and mentored to take future key positions in industry and society. In Norway, basic education is accessible for all. Performance is only emphasised in the final 3 years of secondary education and University.

Dominant cultural Dimensions: Power Distance, Collective versus Individual, Feminine versus Masculine.

Design Problem: The Indian education system struggles with issues such as high cost of providing education, lack of infra-structure, training, governance and supervision teaching staff.

Design Concept: An interactive audio-visual interaction aid to be made accessible to all students in India. The medium of transmission and communication is based on the existing television network.

Mobility of women, Jeddah (Saudi Arabia) versus London (United Kingdom)
Context: Women in Jeddah are not allowed to drive any vehicle nor ride a bicycle by themselves. To get around a woman either has to sit in her husband's or relative's car or walk. She is allowed to take public transport, but the public transport does not offer the same levels of efficiency, comfort, security or status as a private car.
In London, most people choose to go by public transport. The public transport system is among the best in the world, consisting of an extensive network of tubes, buses, trams and trains.

Cultural Explanation: In Saudi Arabia, there are strong conceptions within society concerning gender roles. In addition, Islam is often used as an excuse for the strict limitations to women's rights in Arab countries.

Dominant Cultural Dimensions: Masculine versus Feminine, Power Distance, Uncertainty Avoidance.

Design Problem: How can women in Saudi Arabia be given the possibility to get around in Jeddah on their, own in a comfortable, safe and fashionable way, without violating the strict laws and moral judgements limiting their freedom?

Design Concept: A comfortable, safe and moral proof public transport network of exclusive women's mini buses. The buses are equipped with spyglass windows to prevent men on the outside form seeing the travellers. An automatic payment system is put in place to avoid "indiscrete" interaction.

Making Contact, France versus Norway
Context: In Norway it is normal to pretend you don't see people you don't know when you pass them on the street. Sales representatives pretend they aren't there so that customers will not feel disturbed or embarrassed. Conversation in shops is usually limited to the strictly practical and polite phrases are not widely used. In France smiling, hugging and talking to people you don't know is perfectly normal.

Cultural Explanation: Norwegian culture is more diffuse than French culture, meaning that Norwegians will act reserved towards strangers, and let relatively few people into their "public space" of interaction. However once inside the sphere there is a lesser distinction between close and distant friends. French culture is more specific, meaning that people are more open to strangers but that fewer people will be let into the inner sphere.

Dominant Cultural Dimensions: Masculine versus Feminine, Power Distance, Collective versus Individual.

Design problem: How can Norwegians be helped to make contact with each other without making them feel embarrassed? How to give them an excuse to cross the line of shyness and make contact with strangers?

Design Concept: Based on the mobile phone network, everyone who would like to get to know new people adds a new application on their mobile phone. Each user registers personal information and preferences, e.g. what kind of contact they like to make, on their device. Whenever you come close to another person with matching preferences, both devices will make a little sound or a vibration. Knowing that the device has matched each other interests, you have an excuse to start a conversation.

Waste Collection: Norway versus Singapore
Context: When it comes to efficiency on a macro level, Norway has a slightly better waste collection system than Singapore, but is experienced less convenient by the end-customer. For example, in Singapore the waste is collected everyday and complete sorting is done by workers at a waste plant. In Norway the users have to pay high taxes for the service, which is run by the local authorities

Cultural Explanation: In Singapore immigrant workers do not have citizen rights and minimal wages do not exist. For this reason labour-intensive low-skill's operations like manual waste sorting can be done at a low cost. In Norway garbage collectors are organized in a union, and protected by "anti social dumping " laws. They can leverage a relative high salary, without the pressures to continuously improve their service quality.

Dominant Cultural Dimensions: Power Distance, Long- and Short Term Orientation.

Design problem: How can clients of the Norwegian waste collection system experience a similarly effortless and pleasant situation as those in Singapore? How can collection be made more efficient, while paying waste collectors a decent salary?

Design Concept: A flexible waste collection system, which is based on need. During certain periods and for certain private households, waste containers are not always full. A transmitting indication system integrated in the waste container, which can signal the remaining volume to the garbage collection truck. Upon detection of the signal, a decision can be made to collect or skip the collection. In this way, waste collection trends can be analysed and predicted to optimise its efficiency and savings.

Food retail Shopping Experience, India versus Norway
Context: A large percentage of retail food is imported and sold through grocery retail chains. These retail chains control practically all of the retail stores in Norway. The food retail sector in India is based on traditional grocery shops known as 'Kirana', bazaars, home supplier vendors. They have a good knowledge of the products they sell.

Cultural Explanation: The extended family is a single unit. One of the primary roles of the housewife is feeding the family. Women do most of the shopping and make most food purchasing decisions. Buyer-Seller relationship may remain unchanged from generation to generation and are based more on mutual trust. In Norway, family structures are flexible and nuclear. Social mobility is comparatively higher. According to law, food retailers are responsible to provide formally documented quality products to their customers.

Dominant Cultural Dimensions: Long- and Short Term Orientation, Masculine versus Feminine, Collective versus Individual.

Design Problem: A customer in Norway has most of his food requirements met at a single place. However, expert advice suited to his needs is often lacking, as it is practically impossible to employ a person who is knowledgeable of all products.

Design Concept: A system to capture all necessary details in a user-friendly manner to help the buyer to make purchase decisions. The Camera Culture group at MIT has developed Bokodes that can carry much information, such as Name of Product, Date of Manufacture, Date of Expiry, Contents, Nutrient Value, and Cost etc., and be read by a standard camera.

Elderly Care, Norway versus India

Context: In Norway, senior citizens, who are not able to take care of themselves move to an institution for elderly. In India, elderly will stay traditionally with their oldest son and grand children. While in good condition, they help out with looking after their grandchildren. In case they no longer can take care of themselves, their family will take care of them. Family members, especially their daughter-in law, are responsible for providing physical, medical, financial and emotional support.

Cultural Explanation: In Norway, the society as a whole is responsible for taking care of individuals. Despite this collectivistic thinking, individualism is strong in Norway. Freedom and independence is an ideal for people in need of care as well as their families. In India individualism is less emphasised, making it more likely that extended families will live under one roof. This natural progression in family relationships may limit member's personal freedom and cause conflict.

Dominant Cultural Dimensions: Collective versus Individual.

Design problem: How can elderly people in Norway, who are living on their own or in an elderly home still feel that they are part of their families? How can they maintain contact and not be seen as a burden to their family members and relatives? How can they still participate in the upbringing of their grandchildren?

Design Concept: An interactive easy-to-use photo frame with which you could share pictures with others from a distance. With such an electronic photo frame elderly can keep in touch with their families on a daily basis without feeling like a burden.

The "Design Concepts" were a natural progression of the design problem. However not all solutions could be classified as radical. A bottom-up analysis of the case studies has resulted in the six (6) preliminary categories, which can function as a guide for a broader approach in terms of future external analysis and goal finding. Table 1 classifies the 7 case studies according to the 6 categories.

Referenced to Maslow´s hierarchy of needs [26], the six categories can be classified under the first three level of needs of the pyramid; Physiological, Safety, Love / Belonging. This indicates that a cultural approach towards external analysis and product idea generation in the FEI can be instrumental in the generation of innovative system and or product ideas to improve quality of life and service not only in developing but also developed nations.

Table 1. Classification of case studies according to categories

Categories	Case Studies
Healthcare and elderly care	• Health Care, Denmark versus USA • Elderly Care, Norway versus India
Working and living	• Waste Collection: Norway versus Singapore
Education and manpower development	• Education - India versus Norway
Purchase of food, products and services	• Food retail Shopping Experience, India versus Norway;
Mobility /Transportation of goods and people	• Mobility of women, Jeddah versus London;
Interaction and Communication	• Making Contact, France versus Norway;

7 Discussion and Future Studies

The diversity of case studies, each leading to a proposed design concept, has demonstrated that a cultural and contextual approach towards strategic design should be further explored in the development user interface system and products in the FEI.

As social, political and economical developments in a society are difficult to change overnight, various case studies have illustrated the potential role of design in improving the negative aspects of these developments, usually represented by extreme cultural trends, through innovative design concepts. Potential areas for innovation can be identified by mapping case studies, illustrating extreme trends and developments in certain societies, on a bi-polar scale, supported by Hofstede's "Cultural Dimensions.

Given the potential for a methodology for strategic goal finding based on social, cultural and political differences on a bi-polar scale, future case studies are expected to refine and update the present categories from time to time. However, future case study development needs to be more diverse in context to be able to ascertain that a categorical top-down approach can be applied as a source for external analysis in the generation of innovative system / product ideas, while considering prevalent economic, social and political status quo of their cultures.

At this moment, it is difficult to assess whether design concepts are more innovative by addressing the potential gap of extremities on the cultural bi-polar scale as source for innovation.

References

1. Ehrenfeld, J.R.: Sustainability by Design: A Subversive Strategy for Transforming Our Consumer Culture. Yale University Press, New Haven, CT (2008)
2. Xinyuan, C.: Culture-Based User Interface Design. In: IADIS International Conference on Applied Computing (2005)
3. Marcus, A.: User-interface design, culture, and the future. In: Proceedings of the Working Conference on Advanced Visual Interfaces, New York, NY, USA (2002)
4. Rose, K.: The Development of Culture-Oriented Human Machine Systems: Specification, Analysis and Integration of relevant Intercultural Variables. In: Advances in Human Performance and Cognitive Engineering Research, vol. 4, pp. 61–103. Elsevier, Amsterdam (2004)

5. Cagan, J., Vogel, C.M.: Creating breakthrough products: Innovation from product planning to program approval. Prentice Hall, Upper Saddle River (2002)
6. Drucker, P.F.: The essential Drucker: The best of sixty years of Peter Drucker's ideas on management. Harper Business, New York (2001)
7. Sanders, E.B.-N., Simons, G.: A Social Vision for Value Co-creation in Design. Open Source Business Resource (December 2009): Value Co-Creations, `http://www.osbr.ca/ojs/index.php/osbr/article/view/1012/973`
8. Aryana, B., Boks, C.: Cultural customization of mobile communication devices components Design. In: International design conference - Design 2010, Dubrovnik, Croatia (2010)
9. Buijs, J.A., Valkenburg, R.: Integrale Produktontwikkeling. LEMMA BV. Utrecht, The Netherlands (1996)
10. Belliveau, P., Griffin, A., Somermeyer, S.M.: The PDMA Toolbook for New Product Development. Wiley, Hoboken (2004)
11. von Hippel, E.: The Sources of Innovation. Oxford University Press, New York (1988)
12. Ansoff, H.I.: Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion. Penguin, Harmondsworth (1968)
13. Kim, J., Wilemon, D.: Focusing the fuzzy front-end in new product development. R&D Management 32(4), 269–279 (2002)
14. Kroeber, A.L., Parsons, T.: The Concept of Culture and of Social System. American Sociological Review 23(5), 583 (1958)
15. Hofstede, G., Hofstede, G.J.: Cultures and organizations: software of the mind (Revised and expanded 2nd edn.). McGraw-Hill, New York (2005)
16. Straub, D.W., Loch, W., Aristo, R., Karahanna, E., Strite, M.: Toward a Theory-Based Measurement of Culture. Journal of Global Information Management 10(1), 13–23 (2002)
17. Hofstede, G.: Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations. Sage Publications, Thousand Oaks (2001)
18. Hofstede, G., Bond, M.H.: The Confucius Connection: from Cultural Roots to Economic Growth. Organizational Dynamics 16(4), 5–24 (1988)
19. Massey, A.P., Hung, Y.C., Montoya-Weiss, M., Ramesh, V.: When culture and style aren't about clothes: perceptions of task-technology 'fit' in global virtual teams. In: Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, pp. 207–213. ACM Press, New York (2001)
20. Barber, W., Badre, A.: Culturability, the merging of culture and usability. In: Proceedings of the 4th conference on Human Factors and the Web, Basking Ridge, NJ, USA (1998)
21. Chau, P.Y.K., Cole, M., Massey, A.P., Montoya-Weiss, M., O'Keefe, R.M.: Cultural differences in the online behaviour of consumers". Communications of the ACM 45(10) (2002)
22. Gould, E.W., Zakaria, N., Yusof, S.A.M.: Applying culture to website design: a comparison of Malaysian and US websites. In: Proceedings of the IEEE Professional Communication Society's International Professional Communication Conference and Proceedings of the 18th Annual ACM International Conference on Computer Documentation: Technology and Teamwork, IEEE Educational Activities Department, Piscataway (2000)
23. Kim, W.C., Mauborgne, R.: Blue Ocean Strategy: From Theory to Practice. California Management Review 47(3), 105–121 (2005)
24. Prahalad, C.K., Ramaswamy, V.: Co-opting Customer Competence". In: Harvard Business Review, pp. 79–87 (January–February 2000)
25. Yin, R.K.: Case Study Research: Design and Methods, 3rd edn. Sage Publication Inc., Thousand Oaks (2003)
26. Maslow, A.H.: A Theory of Human Motivation. Psychological Review 50(4), 370–396 (1943)

Socratic Dialogue in Design Education

Neil Matthiessen

University of South Flordia St Petersburg, Program in Graphic Design,
College of Arts & Sciences, 33701, USA
neil@koligdesign.com

Abstract. The Graphic Design Program at the University of South Florida, Saint Petersburg utilizes Socratic Questioning to develop lifelong learners. In today's society lifelong learners continually are re-examining their knowledge and developing their critical thinking skills. The Graphic Design Program teaches the Socratic Questioning methodology so students are able to develop higher-level thinking in their design work. This implementation of higher-level questioning allows the students to critically analyze their ideology of design.

Keywords: Socratic Dialogue, Design, Education, Graphic Design.

1 Introduction

To engage in design there must be an understanding that design is to solve problems, to find new details, to give meaning in something that already has a connotation assigned to its existence. To be able to solve any problem, one must be able to define the problem beforehand. This is where a method, or process, to solve problems needs to be developed as an understanding of the problem. To do this the students need to utilize discovery, develop hypotheses, and know when questions of a fundamental nature are raised. Socratic Method in design allows students to develop a systematic approach in developing new insights and inner understandings. This approach does not provide empirical answers but allows the learner a framework to help them discover how to find the answer. In most situations, Socratic Dialogue can lead to more questions then answers. Although this method does not have a definitive answer, what it does do is develop a way to approach design. This new approach provides order to questions that designers try to answer throughout their careers.

Students, in general at a public university, tend to come from a secondary educational model that does not foster critical thinking skills, but the mastery of facts. In the beginning of the students' design education, students want to find definitive answers, but design thinking does not fit into the widely held belief that statements about the world may be divided into "facts" and "opinion." The students are accustomed to this type of separation and struggle with the ambiguity of design thinking that does not have clear-cut answers. At the same time conventional education does not encourage learners to think for themselves, it encourages them to please authorities. A large majority of students come from this model of education in their primary and secondary schooling. With this as their experience the students have not developed their critical thinking skills and come to the realization that determining the right answer in design is a challenge.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 94–101, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Students in Graphic Design Education

The students that are entering the University of South Florida, Saint Petersburg Graphic Design Program are for the most part traditional students that have either middle or upper-middle class socio-economic backgrounds. The program does have some non-traditional students obtaining a degree for a second career and other cultural and social economic diversity, but the majority of students are from a traditional background. This narrowing of diversity stems from the following, the program is selective only allowing 20 students each year into a cohort system. The factors tend to limit the non-traditional and first generation college student, but being a public institution, these students are more prevalent in the foundations courses prior to the competitive selection process.

For the majority of students that are entering the Graphic Design Program they are dubbed the Millennial Generation. This generation comprises those born between 1982 and 1995 to the late baby boomers. This generation has grown up with technology and knows of nothing prior to the Internet revolution. The students have a close connection to family and friends and have been kept busy with structured activities such as sports, clubs, youth group activates, etcetera. They spend more than half of their waking hours plugged into media of some from; this can be television, the Internet or social media. A large portion of this generation has received weaker K-12 education then previous generations. The students have been so closely connected to friends and families through technology; they have had little experience in self-reflection and self-examination. The students tend to be hyper brand-orientated and measure their status in life through these brands. The students (or parents) view higher education as another service that they pay for and have an entitlement to their degree (Nilson 2010, p. 11-13).

In our society today education is a key component that people are required to have to obtain higher-level jobs. Education has even been tied to national power and the prosperity of a country. As global economics becomes a greater force in the world, so does the need for a nation to have a well-educated society. There are many factors that have pushed western civilization forward, especially in the United States, towards the need for a highly educated work force. One of the main components towards this need for a highly educated work force is the move away from manufacturing to a society more focused in intellectual and technological economics. With this shift of economics and the technological revolution our nation's ability to compete globally has pushed education to the forefront. In the United States this has caused the reexamination of the educational system. New ideologies have been looked at and implemented in this new lifelong learning society.

The policy of lifelong learning has developed into a world issue. International organizations such as the European Commission, the Organization for Economic Co-operation and Development (OECD), and United Nations Educational, Scientific and Cultural Organization (UNESCO) all have developed policies for lifelong learning. These policies that are emerging have played a crucial role in how we educate students today. While the government's agenda pushes lifelong learning into our society, there has been an increase of diversity among the student body from this push - different ages and skill levels, and variances of backgrounds from social to economical. Students also have different reasons for perusing a college degree; some study for interest in the subject, while others seek new skills and vocational education in response to economic changes in their work place (Jarvis 2002, p.159). This has created new challenges in educating the diverse student population. The modes of teaching have changed from a patriarchal teaching style to a student-centered style that instills lifelong learning.

The patriarchal teaching style, where facts are transmitted from teacher to student does little in creating a learning society. The focus is on memorization of information and not learning – which is not just about memorization, but having the ability to be analytical. Learning theorists such as Carl Rogers[1], Malcolm Knowles[2] and Jack Mezirow[3] have all argued that learning is self-actualizing and that what we perceived as traditional teaching is not true learning. Now that the focus in teaching is lifelong learning, this requires the student to be active in his or her own education.

The change in our economical focus as a society has driven this need to develop a lifelong learning society. Technology has also been a driving force behind this need. These changes are attributed to the rapid advancements in technology, the way that we communicate, and even how we socialize. This has created an explosion of information. Colin Griffin states, "New technological knowledge is changing minute by minute and second by second. With this rapid change, it is almost impossible to regard knowledge as a truth statement any longer. We are now talking about something that is relative. It can be changed again as soon as some new discovery is made that forces people to change their thinking" (Jarvis 2002, p. 64). With technology and the rise of social media, information is always being updated and disseminated as soon as a new discovery is made. This rapid cycle of dissemination of information has its good and ills, but it is our new reality. Our students today are a new generation of students that have grown up in this technological society. This requires us to rethink our educational theories and it is these students that are driving that need to develop lifelong learners.

The view of education in colleges has shifted away from an authoritative one to a more consumer based educational system. This is in response to the demands that are being placed on educational intuitions from a variety of directions – which include the need to create an educated work force while meeting the needs of the students' beliefs and attitudes. The public university system faces challenges placed on it from the changing world and it is the belief of some that the traditional educational system is not meeting the needs of our society today. These pressures of change are coming from and are influenced by corporations. Recently John Chambers & CEO of Cisco Systems released a paper calling for fundamental changes in the American educational system. Chambers states in *The Learning Society*, released by CISCO, "Learning is critical to the future of our world and yet our current education systems

---

[1] Carl Rogers (1902–1987) is seen as a major influential American psychologist that was a founder of humanistic approach to psychology. This lead to applications of person-center, and student-centered learning and client-centered therapy.

[2] Malcolm Knowles (1913-1997) was an American Educator and is credited for developing Humanist Learning Theory through the use of learner constructed contracts to guide the learning experience.

[3] Jack Mezirow's work focuses on Transformative Learning Theory, which is the process of leading the learner to re-evaluate past beliefs and experiences and become critically aware of one's own assumptions.

are facing unprecedented challenges. I believe that the effectiveness of how the world achieves genuine lifelong learning is reliant, to a substantial degree, on how well we harness the power of the network to connect and engage learners and educators alike, and to provide access to our collective resources and knowledge. It comes down to scale. Traditional education systems alone, despite the essential role they have played, and will continue to play, in learning, are simply not capable of serving the world's growing and changing needs." (The Learning Society 2010)

Though right in some aspects that traditional education cannot be the only role in the development of a learning society, universities are still the central place where people can learn how to become learners. People need to be taught how to process information in a critical and analytical method for a society to become a true learning society.

The other force that is driving the change in the public higher educational system is the shrinking support from states in budget crunches and increased competition from other online and private universities. This has lead to the university system moving from an authoritative position in the role of education to more of a consumerism perspective. The university system sees students as consumers of education, with a vast array of choices. With the financial constraints placed on public universities they have made concessions to attract and retain students. This idea of the student as the consumer can be felt in the classroom. This shift has changed the dynamics of the role of instructor from a professional authoritarian to one that has to meet the needs of the students at their level. The teaching style has changed based on whatever abilities the students might have. What has happened is that the instructor has to accommodate these abilities of the consumer driven and technological student, but that does not mean that one has to abstain from developing a self-efficient learner.

3 Teaching in Graphic Design Education

In teaching, the most common idea of synonymous teaching is that information, concepts or skills are transmitted to the student by the instructor and improvements in pedagogy are made by means of improving that transmission of information (Strong 1997, p. 61). Though this is a method of teaching, it does not allow the student to become a learner. Education in the 21st century relies on the student taking control of his or her own education, to create a student that is actively engaged in the process in learning.

In design education a lot of the instruction has been hands on, but the method of how to teach the application of design has largely been lecture style. Information is transmitted to the student and the student then applies this information into design work or projects. This is still limiting how the student learns. There is some higher level thinking involved – to take information and apply that knowledge to an artifact, although this does not involve critical thinking. There are many drawbacks to this method of teaching which does not involve the students in any analytical or critical thinking skills.

The failure of lecturing is that it creates a passive environment with no interaction or feedback; it creates a one-way process. In formal lecturing the student plays a minimal part in the learning process and there is little range for reflectivity and experimental connections (Jarvis 2002, p. 55). Though in design education there is room for examination in the form of the artifact, there is still a disconnect that happens from the lecture to the implantation of the artifact. In design there are several issues at hand that the student must be aware of and at USFSP the program is built on the university's mission in creating socially aware and socially engaged students. This type of social engagement requires the students to continually be reflecting on everyday decisions they make. Lectures do not provide this type of engagement. Lectures do not promote discussion or promote critical thinking, they are unable to engage the students' own experience, and are ineffective in teaching value to promote personal and social adjustment, and develop lifelong learners.

The world is shrinking in the global economics that is interlinked by networks that spans geographic, social and cultural groups that are able to interact instantly. Even if a designer is not working globally our own domestic ways of communications have been altered through the technological social revolution. Information is transmitted faster then we could of ever conceived. The designer faces challenges that change on a daily basis. The designer must be able to adapt to those changes. Design is never presented with self-evident truths but rather filled with obscurity and ambiguity and without a learned method to critically assess these truths it is hard to find an effective solution.

To be able to find these truths, and therefore effective solutions, students need to be able to make intellectual judgments. These judgments have to go beyond aesthetic solutions and into social science issues. Design is not just about the artifact it is about communication. The designer needs to understand elements such as linguistics, cognitive psychology, social anthropology, and sociology to develop strong communication. These subject matters tend to be left out of the curriculum and take a second behind aesthetics. To be able to bring social sciences into the design process there needs to be a bridge to understand how to make the most of these concepts.

4 Socratic Dialog in Teaching Graphic Design

In thinking about design we can approach it in terms of philosophy. When teaching students, the students are constantly being asked: What does design mean to them? How does design fit into the larger schematics of everyday events? Can design be used for good? How do you (the student) define design? These and many more questions are profound and cannot be easily answered, and the answer will probably change throughout the career of a designer. Students are given these questions at the beginning of their career and they stay with them when they leave their educational starting point and move into a practicing designer. The answers that the students develop help the student to define himself or herself as a designer. They use these definitions as a guide to develop their self-identity as a designer.

Once a student starts to formulate their definition of what design is, they look at their personal experience with design. The moment that student starts to examine design from the standpoint of their self-experience with design, they begin their examination with a narrow view and are soon in metaphysical darkness. Design is never presented with self-evident truths but rather filled with obscurity and ambiguity and without illumination through a method it is hard to find an effective solution.

This is not to say that students are unable to rationalize vague concepts or that design does not have any answers associated with one solution to the problem. When it comes to design there are numerous solutions to one problem, the problem that most designers face is figuring out what solution works above the rest. When it comes to the students they tend to have a very limited experience in design. With this limited experience the students find difficulty in answering all the questions that must be answered to solve a design solution or coming up with self-definitions for design. Thus, we can determine that students have not explored conceptual thinking at the level that design requires.

To engage in design there must be an understanding that design is to solve problems, how to find new details, give meaning in something that already has a connotation assigned to its existence. To be able to solve any problem, one must be able to define the problem beforehand. This is where a method, or process, to solve these problems needs to be developed as an understanding of the problem. To do this the students need to utilize discovery, develop hypotheses, and know when questions of a fundamental nature are raised. Socratic Dialogue in design allows students to learn to develop a systematic approach in developing new insights and inner understandings. The approach does not provide empirical answers but allows the learner a framework to help them discover how to find the answer. In situations, Socratic Dialogue can lead to more questions then answers. Although this seems to not be an answer for students, what it does do is to develop a way to approach design. This new approach provides an organizational framework for questions that designers try to answer throughout their careers.

In teaching students Socratic Dialogue it is about teaching the principles of the art. The instruction is only a means of guiding the student in how to utilize the method. The knowledge of Socratic education is something that merely cannot be transferred from teacher to student. Knowledge is something that is learned, it is an activity of the mind that requires effort and time. In this method, philosophy is a learned trait; it is not something that is filled with facts that can be recalled at any given moment. In the use of a philosophical method in design it creates a lifelong learner.

The utilization of Socratic Dialogue is one way to develop critical thinking. Students must take responsibility in their own learning by making critical judgments. Through the Socratic method the student develops a new way of thinking, to be able to develop intelligent questions that can lead to intelligent answers. The foundation of this method is not to find new methodologies to the design process, there are better methods to producing new design thinking, it is rather a way to develop insight. The new insight comes from breaking down the walls that we have built up that stops us from understanding a particular knowledge.

Knowledge has many different levels, which encompass all aspects of experience. We can break down knowledge to three distinctive levels. The first level can be described as knowingness. Knowingness is rudimental and is limited. This level encompasses the self, “knowing self” and does not venture from the self. Level two allows us to start to transcend beyond the self. This is where awareness of ones surroundings influences our actions in this world. The third level of knowledge is the ability to understand all dimensions of reality. The third level is what makes us humans; it is the ability to reason. When we create a discourse all three levels of knowledge is present. To be able to truly obtain the third level of knowledge one must break through the other two levels. Fernando Leal states that, “[The] philosophical bottom line, is that deep inside ourselves we have knowledge about the most important things which should concern human beings, namely how we ought to live. This is a peculiar kind of knowledge, calling for a peculiar kind of conversation. But knowledge it is, not just belief or conviction or persuasion of ideology. All these things do exist; but all are barriers to knowledge. They project shadows on our most vital concerns; they obscure the issues; they obstruct the view. In sum, they don’t let us think for ourselves.” (Saran 2004, p. 123)

Practicing graphic design requires all three levels of knowledge. How we navigate through the different levels of knowledge can be difficult. One cannot ignore the self, even if the designer is designing for a subject that does not pertain to the self, the self is subconsciously used as a starting point. We all have experiences and we use these experiences to create an understanding of our world. These self-experiences are also present in the design process. Though, what happens in this process is the intertwining of the self and knowledge and this combination creates the basis for judgments in design. Students use judgments all the time when they start out in design, although they have difficulty discerning between arbitrary forms of knowledge.

Socratic Dialogue creates a method that helps students to reflect on their responses and determine if their response is based on judgments or knowledge. What the attempt to do here is not to use a system that generates new knowledge, but to help students separate and be aware of the different levels of knowledge. This is so the students can begin to separate their “knowing self” from the other levels of knowledge to create informed decisions.

The method of Socratic Dialogue in the design process is slightly modified from the original discourse. Students learn the methodology of Socratic Dialogue in a group to help them develop the process. The goal is to have the students utilize the ideas of Socratic Dialogue to produce private conversations with themselves during the design process. Through these private conservations the students should be questioning their own design assumptions that they have about the subject matter. They should be able to separate their assumptions out of the equation in the design process. For communication to work effectively, the message must be clear to the receiver. Assumptions can lead to unintentional problems in the communication process. The separation of ones own assumptions from the subject facilitates the design process to help distinguish from self-knowledge and factual knowledge. This is an important step for students studying graphic design, it allows them to break down their assumptions and know when to research further into a subject to gain knowledge that will lead to purposeful communication.

The role of the instructor in Socratic Dialogue is to facilitate the discussion. The instructor should avoid being a participator in the process and keep their views out of the group. Students should be the ones to lead the discussion and to develop a conclusion. If the instructor takes control of the discussion, little has happened in the learning process. The main task of the facilitator is to assist the students in the process by keeping the discussion on track. All inquiries and discussion is left to the students and the instructor should allow the discussion to develop through the students’ inquires. The instructor should abstain from directing the discussion or take up any particular stance. When called upon by the students, the instructor can step in to help clarify a topic, but should allow the students to take control of the discussion as soon as it is clarified. The instructor is the mediator in this learning process and is there to insure all students are participating and that no student is dominating or interrupting the process. Socratic Dialogue when applied correctly, allows students to gain control of their learning.

Socratic Dialogue requires everyone to participate and there are few restrictions placed in the process. Participants must be willing to contribute based on experience and not other forms of knowledge. When a participant responds all responses and questions should be genuine. Without genuine contributions the whole process falls apart and getting to any form of knowledge and consensus is lost. Participants should focus on the whole conversation and not on just their own thoughts. Everyone should undertake other participants' ideas and understand them. There are instances that the group needs clarification on an idea presented and it is the responsibility for the facilitator to intervene and create a second question that helps clarifies any misunderstanding. Once the group has reached an understanding then the original question may be taken up again. The query continues until conflicting views prevail, the process does not stop until clarity of the question is reached, or until everyone agrees that an impasse has taken place.

The Socratic Dialogue process for design students works in the same manner, but is utilized in different situations. The process is used for general discussions about graphic design and the role that designers have in today's modern life. With the Graphic Design Program and USFSP focusing on educating students to be socially engaged and responsible it is very important to have students involved in these types of philosophical questions. Another variant in the Socratic process is during critiques. Critiques offer Socratic Dialogue a good venue. Critiques are comprised of talking about the formal aesthetics but when it comes time to move deeper into the meaning and critically looking beyond the aesthetics, the critique can fall apart into random tangents that never address any aspect of the work. A Socratic Dialogue methodology allows for structure to take place during critiques and critical thinking will occur.

Without a strong foundation the solution to a problem will only be as strong as the questions that were contemplated. Socratic Dialogue provides the students with the skills to analyze the problem, to enhance their capacity for independent critical questioning. A student is clouded by their convictions, prejudice and ideologies, which can have an adverse effect on design communication. These obscure the issue but through the use of Socratic Dialogue the students are guided through a methodology that can help their clarification of design theory.

References

[1] The Learning Society (2010), `https://www.cisco.com/web/about/citizenship/socio-economic/docs/LearningSociety_WhitePaper.pdf`

[2] Jarvis, P.: The Theory & Practice of Teaching. Kogan Page, Sterling, VA (2002)

[3] Nilson, L.B.: Teaching at its Best: A Research Based Resource for College Instructors, 3rd edn. Jossey-Bass, San Francisco (2010)

[4] Saran, R., Neisser, B.: Enquiring Minds: Socratic Dialogue in Education. Trentham Books Limited, Sterling, VA (2004)

[5] Strong, M.: The Habit of Thought: From Socratic Seminars to Socratic Practice. Chapel Hill, NC, New View (1997)

Ethnography, Ethnography or Ethnography? What Happens When the Same Word Means Different Things to Different People?

Sasanka Prabhala[1], Daria Loi[1], and Subhashini Ganapathy[2]

[1] Intel Corporation, 2111 NW 25th Avenue, Hillsboro OR 97124, U.S.A.
[2] Intel Corporation, 20270 NW Amberglen Court, Beaverton OR 97006, U.S.A.
{sasanka.v.prabhala,daria.a.loi,subhashini.ganapathy}@intel.com

Abstract. This paper discusses how the notion and practice of ethnography differs for practitioners with different disciplinary backgrounds, especially in a context where ethnography exits academia to enter industry contexts. The paper is divided into four sections. The first provides background to specific experiences and briefly over-views existing literature. In the second part we compare our experiences through an industry case study. The third section proposes a taxonomy, suggesting a number of implications, and providing recommendations on how to integrate cross-disciplinary approaches to expand the scope of conducting user research. The final section wraps up our propositions and provides a number of recommendations.

Keywords: Design, Ethnography, Human Factors, Interface, TV, UX.

1 Introduction and Background

Ethnography is a research strategy used by the social sciences (anthropology and sociology in particular) to gather empirical data on people, society and cultures. Ethnographers typically collect data through participant observation, interviews, and questionnaires and describe the nature of those who are studied through writing. It should be noted that although there are examples where ethnographic strategies are used within market research contexts, when we use the term *ethnography* in this paper we do so within a User Experience research and development context.

Differently from the social science tradition, ethnography today increasingly means different things to different people, depending on their disciplinary backgrounds, since this strategy is increasingly used by many practitioners in varied context and for varied purposes. Design ethnography and human factors engineering provide good examples in this context.

Traditional ethnography processes take from a few weeks to several months, with the researcher observing an environment and noting related practices and cultural traits. As stated by Sommerville, Rodden, Sawyer, Bentley, and Twidale [1], ethnography differs from traditional system analysis by focusing more on the participants and their interactions within the system rather than focusing just on the system, structure and its processing.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 102–110, 2011.
© Springer-Verlag Berlin Heidelberg 2011

A multitude of ethnographic methods have been applied to the technology industry in an ad-hoc manners [2, 3], and studies have shown that applying such methods is highly valuable because of the contextual, qualitative, and rich data they generate. This has been clearly identified by Suchman [4, 5]. Wang and Mainwaring [6] talk about the usefulness of applying exploratory ethnographic methods for understanding Chinese users' experience of online games. Further case studies detail different ethnographic methods applied to air traffic control [7], police database systems [8, 9], and underground railway control systems [10].

Although traditional ethnographic methods are proven to be effective in eliciting user requirements for product, service and system development, an effective inclusion and integration of user experience into requirements engineering processes demands complementary insights and strategies that come from design ethnography and human factors engineering. However, these disciplinary areas apply and see ethnography differently.

In the case of design ethnography, where often the focus is towards developing artifacts and concepts, the researcher utilizes a range of creative elicitation methods to help participants better understand the purpose of the product or service that is being studied and to inspire concept development [11]. While the ethnographer is predominantly interested in understanding human behavior in natural settings without altering it, the design ethnographer is often interested in developing artifacts that support the human interaction in that environment and is typically less troubled if the artifact influences participants' behavior. Designers greatly focus on iterative design and users' needs testing rather than free flow observation of human behavior (12, 13). In brief, Design Ethnography merges – somewhat irreverently – traditional ethnographic approaches with design tools and methods, to gather user data as well as inspirational and design information to directly feed the development of new product, services and systems. Bruseberg and McDonagh-Philp [14], provide an overview of how design methods were used to inform the development of domestic consumer products, identifying the experiences gained in applying focus group techniques to inform the design process.

In Human Factors Engineering (HFE) ethnography aims at understanding the physical, psychological, social, and biological capabilities and limitations of human performance with respect to their environment [15]. It involves detailed observation of society and practices, which are then analyzed and translated to structured needs and usages for many purposes, with technology at the top. HFE is concerned with the design, analysis and development of technology where the primary emphasis is on effectively integrating human interactions with technological products (9) and it employs a number of research methods to understand and develop user models for new products, including observations, interviews, task analysis techniques, Cognitive walkthrough, user interface analysis, and system design techniques [16]. HFE's research methods are usually short term, iterative, and involve technology testing in user setting. Testing results produce rich data (qualitative and quantitative) and can be directly applied to design principles targeted at improving the overall user experience of a product, service or system.

This paper addresses some of the above-mentioned differences in the context of a specific industry-based case study. We discuss how the notion and practice of ethnography differs for practitioners with different disciplinary backgrounds, especially in a context where ethnography exits academia to enter the industry. Through our practice, we have observed a number of commonalities among the way designers, HFEs and more traditional ethnographers (for instance anthropologists) view and practice ethnography. We have observed that these disciplines all focus on understanding people's inclinations, values, behaviors, and identifying gaps towards enabling rich user experiences in different domains. On the other hand, we have however observed a number of differences that offer great opportunities to challenge, discuss and expand what modern ethnography can represent and what it can offer to industry-based research.

2 Experiences

In the next three sections we used a specific case study (MyMedia) to discuss how ethnographic research is conducted from an Ethnographer (section 2.1), Designer (section 2.2), and Human Factors Engineer (section 2.3) perspective. In describing experiences from the three perspectives, we have focused on the following 13 criteria: interview, environment, dependencies, frequency, users, objective and focus, after study, duration, stakeholders, outcome, data type, philosophy, and results. We describe these criteria in section 3, together with our proposed taxonomy.

2.1 The Ethnographer Perspective

As already discussed, ethnography is a research strategy used by the social sciences to gather empirical data on people, society and cultures. Data is collected through tools such as participant observation, interviews, and questionnaires and the nature of those who are studied is described through writing. Traditional ethnography processes take from a few weeks to several months, with the researcher observing an environment and noting related practices and cultural traits.

In 2007 one of the authors conducted an ethnographically informed qualitative study focused on gaining a deep understanding of home-care practices. The study, focused on Sweden, China and Indonesia, included face-to-face interviews with participants selected broadly across the company's segmentation, household and contextual observations, expert interviews, the use of Cultural Probes [17] and to some extents Playful Triggers [18, 19]. Field activities lasted 3 weeks per location and were driven by an observational philosophy.

Before the study the two key researchers (an anthropologist and a designer) developed semi-structured interview scripts (translated into local languages to facilitate mother-tongue research assistants). The planning included audio and video recording of each interview, participant-facilitated home-tours, image documentation of contextual data (for instance, retail shopping), and the design and development of picture diary probes and some group triggers.

Within a context where the intent was to follow key areas highlighted in the interview script, both researchers knew upfront about the likelihood of having to modify focus areas or questions following the flow of each interview. The script was used as a general compass to ensure key areas would be covered but each interview was conducted rather differently, in light of context, interviewee's responsiveness, amount of present household members, unpredicted interferences and so on.

At the end of all the planned field work in the 3 locations, the study generated extensive and rich documentation (in the form of video data, still images, transcripts, researcher-notes and user-generated annotations with imagery) and identified a number of matters related to home practices and a range of related topics. However, as the study was conducted within a corporate setting, where timeframes rarely enable long-term research analysis, researchers had an urgency to quickly cut through the data to provide useful directions to planning and technological counterparts in the organization. As key insights on storage practices strongly emerged from the data and many varied stakeholders had a genuine interest in the topic, researchers focused their efforts in cutting through the data to tell a compelling story on the subject [20]. Although specifically focused on storage practices, the report portrayed a holistic story on the matter, highlighting qualitative aspects and focusing on the depth of people's lives. However, it should be pointed out that researchers, once the storage practice report was completed, kept (and still are) drawing from the data for other projects. In a way it could be argued that from the ethnographer's perspective the there were no dependencies with product development timeline, with limitless future opportunities of harvest insights from different points of view and for different purposes and projects. Researchers decided in this case not to have shadowing activities around the participants (although it is standard practice to do so in these types of endeavors).

To summarize the key points that distinguish in our view how traditional ethnography approaches the field within industry contexts, we can say that it collects *qualitative data* through an *observational* philosophy focused on *understanding* a wide *variety of users* in their own contexts during *medium/long* and *open-ended periods* of time, through *semi-structured and semi-scripted* methods *open to evolving events* and to *shadowing techniques*, to generate for a *wide variety of stakeholders holistic and abundant insights* around the *depth of people's lives*.

2.2 The Designer Perspective

As mentioned in section 1, Design Ethnography merges – somewhat irreverently – traditional ethnographic approaches with design tools and methods, to gather user data as well as inspirational and design information to directly feed the development of new product, services and systems. In this section we discuss how the designer that co-conducted the ethnographic study described in the previous section re-utilized and re-conceptualized key insights from the research.

During and after fieldwork, while sharing with the anthropologist an observational approach focused on qualitative data gathering, the designer was clearly open to and interested in influencing participants with design tools and prompts, was generally operating using a fluid and often semi-improvised research structure equally aimed at general or specific understandings (e.g. usability) and equally flexible to undertake research in any type of context (e.g. lab or home). The type of participants that interested the designer varied depending on the phase of the project, timelines were equally flexible and case-by-case (contracting between short and medium periods) and the overall approach was focused on generating and representing rich design data (specific, holistic and inspirational) for engineering, designing, and marketing stakeholders. Throughout the research the designer was result-focused although it should be noticed that the approach generated more of the needed data, which was subsequently used for design and concept development activities.

After field work, Whilst the insights and recommendations developed through the study deeply influenced the thinking around Media Aggregation architecture, the designer was eager to look for opportunities to apply and further probe findings. At that time two important developments were under way: an increasing ecosystem focus on storage and media access on TV and Intel's collaboration with Yahoo! Inc. "to provide a full-featured software framework named Widget Channel that allows TV viewers to enjoy rich Internet applications called TV Widgets while watching their favorite programs" [21].

In light of these developments, the designer worked with various stakeholders to drive the design and development of *MyMedia*: a media aggregation widget focused on people's practices around media consumption, collection and navigation and aimed at providing access to personal media (located on any connected device) from the TV (Fig. 1). The widget, iterated through User Experience Assessment (refer to section 2.3), was launched at the Consumer Electronics Show (CES) in 2009.

During the development of this widget a number of restrictions did not enable the team to develop all ideas around TV-based media navigation and consumption that emerged from the field work. Additionally, user research was showing that 3D User Interfaces (UIs) represented an opportunity to address users' expectations around TV UIs [22]. MyMedia 3D was therefore developed in early 2009, to showcase a number of concepts related to personal media (3D navigation, consumption, sharing, contextual recommendations and promotions).

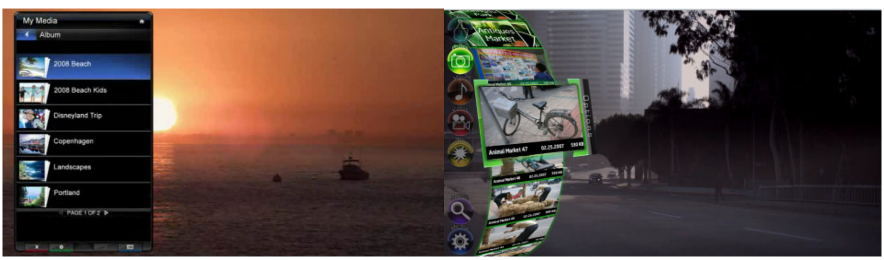


Fig. 1. MyMedia Widget: Widget Channel version (left) and 3D version (right)

2.3 The Human Factors Engineer Perspective

As already mentioned, Human Factors practitioners employ different tools/methodology such as usability testing, user experience (UX) testing, and ergonomics assessment to understand the capabilities and limitations of human performance and in applying the understandings to the design, development and deployment of systems and services.

This section describes through a Human Factors Engineer view of ethnographic research using the MyMedia 3D case study. The objective/focus of this research was to identify MyMedia 3D key value propositions by understanding users' likes, dislikes, perceptions, reactions to the product, initial responses to, and interest in, its features and capabilities. Since the objective/focus of MyMedia 3D research was on usage models and user experience, user experience testing methodology was a natural choice for the HFE.

The type of users targeted for user experience testing usually varies and depends upon: population size for which the product/concept is being targeted; when during the product development testing needs to be carried out; and whether users are technologically savvy or not. For this study, the researcher: screened and selected 24 technology savvy users; organized a controlled usability lab mimicking a living room; developed a script enabling users to (1) interact with MyMedia 3D in short focus groups and one on one interviews, (2) measure their experience and (3) produce rich yet specific data points (quantitative and qualitative) that would address the objective/focus of the research; targeted engineering and the design teams as key stakeholders. It should be noted that this approach to ethnography is structured and does not allow room for impromptu changes to methodology and testing procedure.

Two User Experience studies were conducted in 2009 to test and iterate the design of MyMedia 3D. In both sessions, users' response was overall positive and encouraging, with users providing specific emphasis on:

- 3D look – for its *cool factor* and *smoothness*;
- opportunity to visually view many elements at once;
- possibility of sharing media;
- possibility to get related content, especially in relation to music; and
- live TV as picture-in-picture during non-immersive activities such as browsing lists of photos.

In the first UXA round however a small subset of participants expressed they felt overwhelmed with the graphic layout of the media wall and its animations. This prompted us to reiterate and modify the design, with a resulting positive progression in the overall response to 3D interfaces, depicted during the second UXA test as "cool" yet balanced by usefulness and subtler effects.

3 Taxonomy

To summarize the key points that distinguishes in our view how traditional ethnography approaches the field within industry contexts, we can say that it collects *qualitative data* through an *observational* philosophy focused on *understanding* a wide *variety of users* in their own contexts during *medium/long and open-ended periods* of time, through *semi-structured and semi-scripted* methods *open to evolving events* and to *shadowing techniques*, to generate for a *wide variety of stakeholders holistic and abundant insights* around the *depth of people's lives*.

Similarly, from a designer view of ethnography, we can say that it collects *qualitative* data through *observational and influencing* philosophy focused on *understanding users (case by case dependent)* during *short/medium/long* (case by case dependent) periods of time, through a *fluid* method, to generate <u>*rich*</u> *design data (often more than wanted)* to the *engineering, designing, and marketing stakeholders*.

Likewise, from a Human Factors Engineer view of ethnography, we can say that it collects *quantitative and qualitative* data through *observational* philosophy focused on *understanding technologically savvy users* interactions during *short* (and linked to *product development timeline*) periods of time, through *structured and scripted* methods, to generate *specific data points* to the *engineering and design stakeholders*.

Table 1 proposes a taxonomy that compares and contrasts how ethnography is perceived across the three disciplines. Table 1 also highlights how to integrate cross-disciplinary approaches to expand the scope of conducting user research.

Table 1. Taxonomy: HFE/Designer/Ethnographer views of ethnography

Criteria	**Ethnographer**	**Designer**	**Human Factors Engineer**
Research philosophy	Observe (no influence)	Observe (with influence)	Observe
Research structure	Semi-open	Fluid	Closed
	Open to evolving events	Semi-improvisation	Planned/controlled
Objective & focus	Understanding	From UX/Usability to general understanding	Usage models, User Experience, Usability
Interview style	Semi-scripted	Fluid	Scripted
Target Users	Anyone	Anyone, often savvy	Savvy on topic
Key Stakeholders	Many stakeholders	Eng., design, marketing	Engineering and design
Research location	In-home and context	Anywhere	Controlled (lab…)
Research Duration	Long	Flexible	Short
Dependencies	No, open-ended	Case by case	Linked to prod. dev.
Data focus	Depth on people's lives	Design richness (specific, holistic and inspirational)	Get specific data
Data type	Holistic	In-between	Specific
	Qualitative	Qualitative	Quantitative/qualitative
Results	Gets lot of data and some might not be of specific/immediate use	Gets what looking for but accumulates extra data	Always gets what was looking for
After study research	Shadowing	Often no, but open to it	No after contact

4 Propositions and Recommendations

This paper through MyMedia case study discussed how ethnography was conducted by practitioners of different disciplines (Ethnographer, Designer, and Human Factors Engineer) in an industry setting. Although, each discipline focuses on understanding the user needs, desires, and requirements, they are done very differently with both pros and cons. An integrated approach that combines these different methods from different disciplines offers complementary value to the user research community. This requires each discipline practitioners to understand the tools in the other two disciplines – partly can be obtained from training in an industry and partly via experience. To support an integrated approach to user research, it is critical to have the support of the following propositions.

- Meta-level: It is imperative to design a meta-level method and approach upfront (before planning and commissioning a new project or research) to incorporate in a balanced way the different disciplines point-of-view to enable full iterations across the board.
- Organization resource and commitment: To enable meta-level method and approach upfront, it is essential for an organization to resource, fund, and manage the three disciplinary teams properly.
- Manager support: Also requires capable managers to properly support and nurture the cross-disciplinary teams within the organization. Also critical is to support and help individuals in the teams in their career development within and outside the organization.
- Business group alignment: The likelihood of these disciplinary teams to succeed is higher if the teams are embedded within a business group. This would enable the teams to interact and absorb the business group's demand and needs while being fully immersed at the same time being detached conducting research that addresses the business group's needs.
- Tailor findings: Communicate and share research findings with key stakeholders that combines all the goodness from each of the three disciplines (rather than focusing on own discipline's findings) to the business group for greater influence and impact.
- University training: Exposing early on, the new workforce to other disciplines methodologies and approaches would help in understanding the nuances between the disciplines and more importantly how to work in a collaborative way with other disciplines.

Acknowledgments. We wish to acknowledge all colleagues at Intel Corporation that contributed in different ways to the reported case study.

References

1. Sommerville, I., Rodden, T., Sawyer, P., Bentley, R., Twidale, M.: Integrating ethnography into the requirements engineering process. In: Proceedings of IEEE International Symposium on Requirements Engineering, pp. 165–173 (1993)
2. Madanmohan, T.R., Navelkar, S.: Roles and knowledge management in online technology communities: an ethnography study. International Journal of Web Based Communities 1, 71–89 (2004)
3. O'Brien, J., Rodden, T., Rouncefield, M., Hughes, J.: At Home with the Technology:An Ethnographic Study of Set-Top-Box Trail. ACM Transactions on Computer-Human Interaction 6(3), 282–308 (1999)
4. Suchman, L.A.: Office procedures as practical action: Models of work and system design. ACM Transactions on Office Information Systems 1(4), 320–328 (1983)
5. Suchman, L.: Plans and Situated Actions. Cambridge University Press, Cambridge (1987)
6. Wang, Y., Mainwaring, S.D.: Ethnography at play: An exploratory case study of Chinese user's experience in online games. In: Proceedings of the CHI 2008 workshop on evaluating user experiences in games, Florence, Italy (2008)

7. Harper, R., Hughes, J., Shapiro, D.: Harmonious working and CSCW: Computer technology and air traffic control. In: Benford, S.D., Bowers, J.M. (eds.) Studies in Computer Supported Cooperative Work Theory, Practice and Design. North-Holland, Amsterdam (1991)
8. Ackroyd, S., Harper, R., Shapiro, D., Hughes, J.: Information Technology and Practical Police Work. Open University Press, Milton Keynes (1992)
9. Maguire, M.: Methods to support human-centered design. International Journal of Human-Computer studies 55(4), 587–634 (2001)
10. Phillips, C.A., Repperger, D.W., Reynolds, D.B.: Human Factors Engineering. Wiley Press, New York (2000)
11. Salvador, T., Bell, G., Anderson, K.: Design Ethnography. Design Management Journal 10(4), 35–41 (1999)
12. Schuler, D., Namioka, A.: Participatory design: Principles and practices. Lawrence Erlbaum Associates, Hillsdale (1993)
13. Forlizzi, J., Battarbee, K.: Understanding experience in interactive systems. In: Proceedings of the 5th conference on Designing Interactive Systems, Processes, Practices, Methods and Techniques, pp. 261–288 (2004)
14. Bruseberg, A., McDonagh-Philp, D.: New product development by eliciting user experience and aspirations. International Journal of Human-Computer Studies 55(4), 435–452 (2001)
15. Human Factors and Ergonomics Society, `http://www.hfes.org/`
16. Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C., Jenkins, D.P.: Human Factors Methods: A Practical Guide for Engineering and Design. Ashgate Publishing, Aldershot (2005)
17. Gaver, B., Dunne, T., Pacenti, E.: Cultural Probes. Interactions, 6(1), 21–29 (1999)
18. Loi, D.: The Book of Probes, in Lavoretti per Bimbi–Playful Triggers as keys to foster collaborative practices and workspaces where people learn, wonder and play. PhD Thesis, RMIT University, Melbourne Australia (2005)
19. Loi, D.: Reflective Probes, Primitive Probes and Playful Triggers. In: EPIC07–Ethnographic Praxis in Industry Conference, Keystone, CO (October 3-6, 2007)
20. Loi, D.: Digital home storage beyond devices: a cross cultural perspective on how people keep, protect and find what they value. Keynote, UPA Europe–European Usability Professionals Association Conference, Turin Italy (December 4-6, 2008)
21. Intel Corporation: Consumer Electronics – Widget Channel Overview, `http://www.intelconsumerelectronics.com/Consumer-Electronics-3.0/Widget-Channel-Overview.aspx`
22. Loi, D.: Changing the TV industry through user experience design. In: 14th International Conference on Human-Computer Interaction, Orlando, Florida (July 9-14, 2011)

Improved Usability through Internationalization

Carsten Witte

Siemens AG, Healthcare Sector
Magnetic Resonance Division,
Allee am Röthelheimpark 2, 91052 Erlangen, Germany
carsten.witte@siemens.com

Abstract. This article is a field report about usability and internationalization based on the example of Japan. It will explain, based on examples, how proper internationalization improves the usability of software.

Keywords: software, internationalization, I18N, localization, L10N, globalization, G11N, Japan, Japanization, J10N, language, usability.

1 Introduction

Over the last years Siemens and many other big software and hardware vendors in the life-science industry have focused more and more on workflow in order to support the customer. Solutions weren't centered on single, selective tasks anymore, but on complete chains of tasks. The process of scanning medical images, processing and reading them and finally writing a report of the findings is now seen as one holistic clinical workflow, instead of the former single workstations that didn't know anything about the next step.

I have been working for the Magnetic Resonance division of Siemens Healthcare since 2001. I also have an extensive background in user-interface design. Five years ago, I had the opportunity to visit Japan several times to analyze and learn about the differences between Japanese, American and European methods of operation. At that time, I was responsible for the native language support of our software, too. In a meeting with the then-chairman of the Regional Unit Japan, Bernd Rennebaum, we spoke about these topics: the customer's workflow, supporting Japanese in our software and how to reach an even higher level of usability. And at this meeting, he coined one sentence that ever since became sort of a mantra for my daily work:

> *"How can we better support the workflow of our customers than providing the software in their own language?"*
> *Bernd Rennebaum, Siemens AG, 2006*

This focuses on various issues of usability that need to be taken into account: readability, accessibility, comprehensibility and a continuous flow of reading over all devices in daily life.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 111–116, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Why Localize?

Similar to Germany, in Japan, English is the first foreign language taught in every school. Although some Japanese hesitate to actually speak English, the majority read and writes it fluently. So, why should we translate our software at all?

> *"If I am selling to you, I speak your language.*
> *If I am buying, dann müssen Sie Deutsch sprechen."*
> *Willy Brand, 1913-1992, German Chancellor*

Japan is the one of the biggest markets for magnetic resonance devices, with almost the same size as the whole of Europe combined. So, a healthcare-provider will want its share of that market. And when you enter a clinic in Japan, you're of course greeted in Japanese, both spoken and written. All patient-information is in Japanese. Every workstation speaks Japanese: from the registration desk, over to the radiology counter, the scheduler, and even the filming device. Now imagine your device in that consistent Japanese world… displaying English only. Soon this device will be considered an outsider, an obstacle in an otherwise seamless workflow. This is something you really want to avoid. You will want to allow the user to stay in one language, their native language, without always having to translate in the back of their heads, without recalculating calendars and measurements. You will want to enable the users to read their language and use their keyboards. Aside from all valid business figures, you will want to…

> *"Show respect for you customer's linguistic and cultural needs."*
> *Bill Hall, MLM Associates, Inc., 2006*

By increasing the usability of software, you're increasing the customer's satisfaction, the loyalty of the customer base and ultimately your sales figures.

2.1 How to Localize?

Internationalization and localization as part of globalization are both easy and very, very complicated. It's easy, because you simply have to follow a basic set of rules in order to ready your software for the international market - especially if you start your project in today's software-landscape based on Unicode. Older systems actually had to deal with a huge load of code-pages, font-problems and conversion issues.

First and foremost you have to separate the code itself from all user-visible text throughout the whole software to enable two colleagues to work independently on the project without too many interference: the software-developer and the translator. Both get their set of files and you can compile and translate independently.

This way, you have to compile the project only once and create the product by combining one constant code-base with the language of choice. This approach saves costs all along the product cycle: development, integration, test, localization, and shipping.

At Siemens Healthcare MR, we provide the user-interface in six standard-languages: English, German, French, Spanish, Chinese, and Japanese; and the user-documentation in up to 30 languages.

2.2 What Is So Special about It?

Second, you have to make sure that everything cultural isn't hardcoded anymore either. This includes grammar, calendars, measurement systems, and in rare cases even the color scheme. Below, I will illustrate some major issues:

Context. Software developers are trained to break up larger, complex modules into smaller functional packages. Sometimes this splits sentences into multiple parts, too. When, later on, a translator sees these parts without knowing the context, he might translate the wrong meaning. For example "back" can mean going back to the previous page, or it could mean the rear part of the human body. Mixing these up, the result might be a good laugh if you're lucky or a misdiagnosis if you're not. The solution is to keep sentences together and give the translator a chance to do their work correctly.

Grammar. Although some languages share rules on how to order words of a sentence, or how a plural is created, there are even more substantial differences. Therefore it is unlikely, that you can come up with a one-size-fits-all language-independent grammar-creator. You have to actually implement for the target language. For example Japanese ordinal-postfixes vary not only on the number, but also on the size and shape of the counted object. There are languages, where the plural ending is based on whether the quantity itself is even or odd. German features five cases, Russian six.

In order to address these issues early, anchor the target languages into the requirements of your software and involve the translation team very early along the development process.

Numbers. While in English and German the order and listing of numerations is nearly identical, Japanese differs completely. English list the partial number first: "7 of 42" Japanese starts with the whole number: 「 」 While English places words in front of the numbers of ranges: "from 7 to 42" Japanese features postfixes 「 」. So do not try to arrange the words in the software, but create placeholders which can be reordered as needed by the translator.

Sort Order. Sort order and comparison are very language dependent. While German is sorted by dictionary or by phone-book, Spanish knows a traditional and a modern sort-order. While Taiwan sorts primarily by stroke-count, mainland China prefers pronunciation. Japan then sorts by either XJIS or by Unicode, and sometimes by radical. But there is a good chance that the operating system (Windows, Linux, Mac-OS) already features complex sort-order-algorithms, so you have a good chance to be on the safe side by simply not trying to roll your own.

On that note: today most operating-systems have undergone an extensive test in all sorts of target languages and mostly fulfill not only language, but also even legal requirements of many countries, for example GB18030 for China. It improves the consistency of usability not to deviate from these standards too much. Stick with the common user-interface. And of course, reusing existing, tested controls will reduce development costs dramatically.

Names. People do not only cherish their own name, it is of vital importance to recognize a person correctly, especially if it comes to medical treatment. Because of this importance, the software shall display the name precisely all the time, taking into account the local differences for various countries.

Japanese names can have three equivalent name-representations in parallel. These are written in three distinct writing systems: kanji, kana—katakana or hiragana—which are a phonetic rendering of the kanji, and in a romanization, called romanji. Since there is not one-to-one relationship between kana and kanji, similar kana may represent different names. Therefore, you need to display all available name-representations when applicable. For example, the name "Miki" written in katakana as " キ" has about 70 distinct kanji-representations.

Calendars. Different calendars and date-representation in general, is another topic already covered by most operating-systems. Not every country is based on the Gregorian calendar. In Japan the emperor-calendar is used, especially for official documents. In that calendar "November, 29th 1966" is represented by「 和41年11 29 」.

And there are many, many more issues to look out for: the right-to-left reading order for Arabic and Hebrew or compound letters in the Indic derivates. Parsing strings can be difficult when it comes to different white-spaces and abbreviations. Parsing numbers might be impossible, because a character might only look like a number, but actually it isn't one. Or it doesn't look like one at all, but actually is.

Re-casing of words might mangle the meaning. The most prominent example being: élève → Elève → ELEVE→ eleve. Sometimes different Unicode points are depicting the very same character, almost not distinguishable to the human eye, thereby threatening the safety of internet-addresses, for example. Usability also means helping the user to avoid errors and enabling them to use the software without having to worry all the time.

3 Usability Testing

One of the most important tasks is to test and verify functionality and translation with the target-user-group. The issue with the Japanese names for example came up at a usability-session with customers at large university-hospitals in Tokyo. We discussed the problem, its background and customer wishes, provided prototypes of possible solutions and the best fitting version for both customer and vendor was chosen. This procedure achieved a technical feasible, cost-effective solution that really solved the problem at hand.

When you compare the following two screenshots, you will recognize that they not simply translated into their target-languages, but also handle different number- and measurements-systems. Yet, the key-feature is the sub-dialog in the Japanese version that enables the input of the three name-representations.

Both dialogs provide specialized usability to different groups of users, and these groups only, making localization an enabler to usability. If mixed up, the English version would lack important functionality for Japan and the Japanese version would make entering names in England very, very complicated.

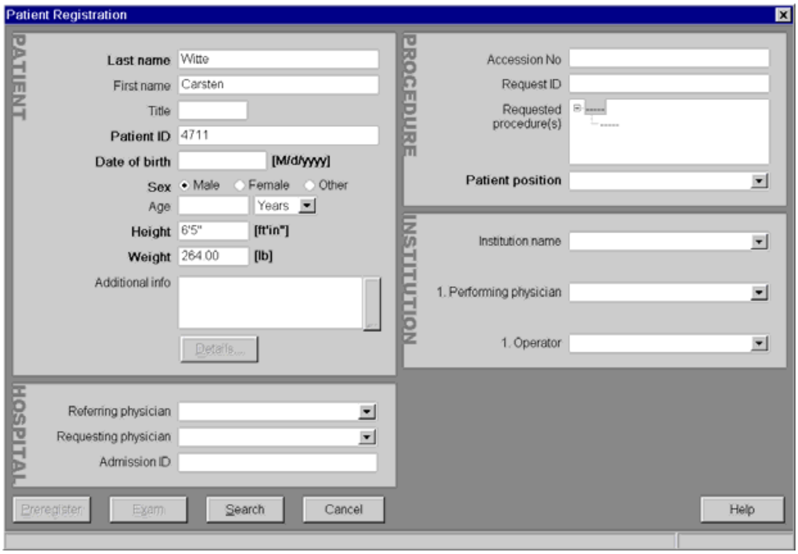


Fig. 1. English dialog for patient registration

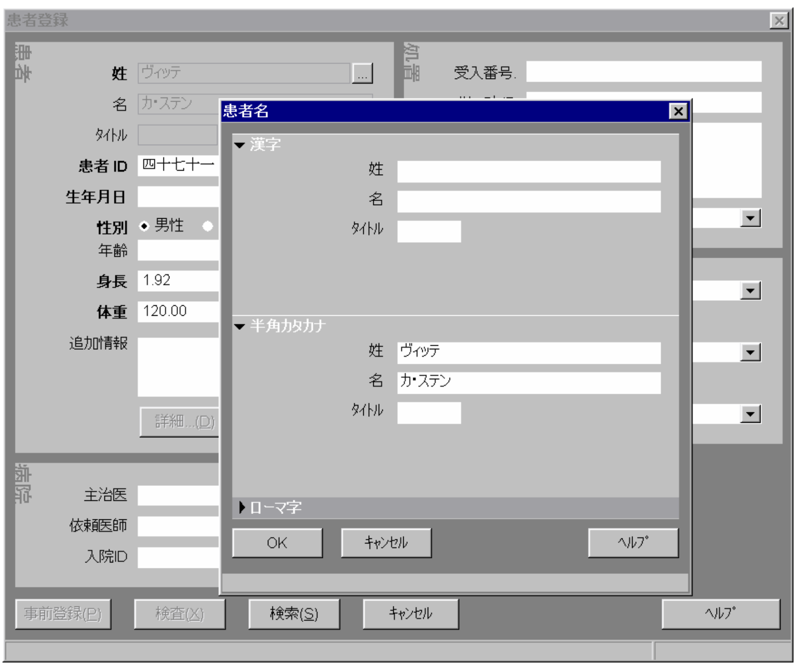


Fig. 2. Japanese dialog for patient registration

There are many tools and techniques to constantly test the localization of your product. Especially the screen-layout should be monitored regularly, because it might break when text is translated from a shorter language like English to a much wider language like German that on average takes up to 30% more space. Give the actual user a chance to comment on used fonts and overall readability.

4 Summary

Although at first glance this sounds like a lot of difficulties and never-ending work, over 80% of these issues are easily solved, once you are aware of the problem. For each technical complication, there is a big chance that the solution is already out there.

Your biggest task is to make sure that everyone involved knows that these efforts are done for the good of the customer. Because so many people involved in the creation of software never meet the customer, it is your job to communicate and motivate the customer's right to get localized, easy-to-use software.

If you want to market your software with usability, make sure that the software is properly localized, too.

ISO 20282: Is a Practical Standard for the Usability of Consumer Products Possible?

Nigel Bevan[1] and Simon Raistrick[2]

[1] Professional Usability Services, 12 King Edwards Gardens, London W3 9RG, UK
[2] Foviance, 14 Bonhill Street, London, EC2A 4BX, UK
mail@nigelbevan.com, simon.raistrick@foviance.com

Abstract. The current ISO 20282 standard is intended to provide a test method that can be used to assess the ease of operation of walk-up-and-use products and consumer products. The standard is currently being revised to improve the cost-effectiveness of the test method, to broaden the scope to include aspects of usability that go beyond ease of operation, and to more clearly define how to obtain reliable results. But the challenge remains to provide a standard that will be cost-effective and useful to manufacturers, purchasers and test houses.

Keywords: standards, usability, consumer products, measurement, test method.

1 Why Aren't Consumer Products Usable?

There are many reasons why a product may not be usable. Some manufacturers try to make products usable but lack the skills or measures to do so, whereas others consciously decide not to adopt usability best practise. This can be due to internal politics, or due to a customer experience management strategy that prioritizes other aspects of the customer experience over usability. However, one common factor across many industries is that senior executives are often unaware that the products their companies produce are unusable, due to a lack of effective comparative measures.

In contrast, the market is deeply concerned with usability, as is often evidenced by poor results in product reviews. For example, the ViaMichelin X-930 portable GPS navigator received a very poor review from PC Magazine [6]:

> "Overall, I was sorely disappointed with the ViaMichelin X-930. Its menus were difficult to navigate and inefficiently laid out. For example, to find a list of the nearest gas stations from the map view takes a time-consuming ten screen taps. … For entry-level GPS devices, there are others on the market that are better and easier to use."

Almost all consumer organisations and magazine product reviews rate usability as an important factor, so why do they not use more thorough and comparable usability measures, and why do manufacturers not give usability a higher priority?

Unfortunately people have to purchase many products without the opportunity to try using them, and shop assistants only have very limited knowledge about usability [3]. A UK TV program recently highlighted customers' frustration at not being able to try mobile phones before purchase. This led to a redesign of a store to allow more interaction with products before purchase [14].

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 119–127, 2011.
© Springer-Verlag Berlin Heidelberg 2011

But even when it is possible to try products in a shop, it is difficult to assess the usability because usability can only be fully experienced by using a product in real life. And this is not feasible within the short time frame of the purchasing decision. It is a challenge to systematically identify the important tasks and try them all. At best, it is usually only possible to practice some basic tasks with a device [11].

In general, there seems to be little incentive for a manufacturer to worry about whether consumers can use their products. Indeed, the early adopters of innovative technology are more concerned with the excitement and status of owning and using an innovative device, than with whether it is actually useful.

As far as the market is concerned, this is often true of technologies that are in their infancy, but for users of technologies which have gone through several iterations, lack of usability is a time-consuming frustration which gradually moves higher up their agenda. Some consumers return products they have purchased complaining they are faulty, when the real explanation is that they were unable to work out how to use the product. For example, a study of mobile device returns in the United Kingdom showed that 1 in 7 mobile phones was returned within the first year of purchase as faulty [12]. Of these returns, about 63% had no hardware or software fault but the reported problems related to usability, mismatch with user's expectations or issues relating to the configuration of the handset.

It would be a great benefit if reliable, comparable information about usability could be provided to potential purchasers. This was the original objective of the ISO 20282 standard.

2 Standards for Usability

The ultimate test of the usability of a product is that users are effective, efficient and satisfied when using it [1]. These criteria were established in ISO 9241-11 in 1998, and subsequently formed the basis for the Common Industry Format for Usability Test Reports [3], which became ISO/IEC 25062. The intended audience for this standard was corporate purchasers, providing a standard format for suppliers to provide usability information to potential purchasers.

In 2000, work started on a proposed standard for assessing consumer product usability [2]. Gradually the work focused on walk-up-and-use products (such as ticket machines) as these were identified as being easier to test reliably as they have well defined user groups and tasks.

The standard ISO TS (Technical Specification) 20282-2 published in 2006 specifies in detail the procedure for summative testing intended to ensure that results are reliable and consistent.

To cover consumer products, two further standards in the 20282 series were developed, ISO PAS 20282-3 Test method for consumer products, and ISO PAS 20282-4 Test method for the installation of consumer products. These were published with the lower status of Publicly Available Specifications to reflect the need for the validity of the methods for consumer products to be checked.

3 User Requirements for Measuring Consumer Product Usability

With little evidence emerging of the practical use of ISO 20282, in 2009 a new ISO group was established [8] to review the approaches to assessing consumer product usability.

The potential benefits of providing a standard summative test method for different stakeholder groups was reviewed:

a) For corporate and government purchasers of walk-up-and-use vending machines, usability is essential for their success. Information about usability could contribute to purchasing decisions, support a more effective tendering process, and improve user adoption.
b) Manufacturers of both walk-up-and-use and consumer products could benefit from using usability test results in marketing to differentiate their products as well as in quality control.
c) Organizations commissioning or carrying out tests (typically consumer organizations) could benefit from a valid test that provided more reliable results, although it might be more expensive than current test methods.
d) Consumers would have objective, comparable information about the level of usability of a product at the pre-purchase stage.

The test method has to be cost-effective to implement, and there was some concern that the test method in 20282-2 may be viewed as unacceptably expensive because of the large number of participants needed to obtain statistical significance. In order to estimate a population success rate greater than 75% with 95% confidence, which was seen as a suitable success rate, a maximum of one unsuccessful user out of 17 would be tolerable, requiring a test of a minimum of 17 users.

Investigation showed that the most recent statistical procedures (the adjusted Wald test [13]) required fewer participants, and the less stringent criterion of an estimated population success rate greater than 75% with 80% confidence could be achieved with one unsuccessful user out of 9. This would bring the test method in line with common practice of testing 10-12 participants to assess usability.

4 Summative Test Procedure

The test procedure proposed is specified in sufficient detail to ensure reliable and consistent results. (The revised version of 20282-2 has made few changes.) The steps are:

a) Ensure that the people who will carry out the test have an acceptable level of knowledge in usability testing.
b) Identify the product to be tested.
c) Identify whether the product is within the scope of the standard.
d) Identify the main goals of use of the product.
e) Define the main goals to be tested.
f) Establish criteria for main goal achievement.
g) Identify the tasks.
h) Specify the user groups to be used for testing.

i) Specify relevant environmental characteristics which affect usability.
j) Check that the product is compatible with intended user characteristics.
k) Decide whether to test one or more user groups.
l) Identify which measures are required, and whether there is a particular value which defines a success criteria for each measure, or whether it is a purely comparative measure.
m) Decide the desired confidence level for the testing.
n) Specify test scenarios and conditions, with clear criteria for goal achievement.
o) Recruit a representative sample of users (that represents the intended user group(s) of the product).
p) Test the product in an environment that resembles as closely as possible the environment in which the product would be used.
q) Establish a written test procedure using best practice usability testing, for example ensuring that the user is not led or prompted.
r) Measure success rate, task time and satisfaction.
s) Calculate effectiveness (percentage success rate), efficiency (median task time) and satisfaction (mean questionnaire scores).
t) If success criteria for task time and satisfaction were identified, the proportion of users that met these should be calculated.
u) Prepare a full report and, if required, a short summary.

5 Reliability of Usability Measures

As part of the development of ISO20282-3, a detailed analysis was made of the factors that would determine the reliability of the test method for a range of products and systems used by consumers. The conclusion was that it would be difficult to obtain reliable measures of usability for certain types of “complex” products where any of the following are true:

- It is not possible to define the user's goals in a clear and repeatable way (for example where these goals vary considerably between users).
- The criteria for success are difficult to define (for example where they include an element of subjectivity or creativity).
- It is not possible to reliably measure the success of the outcome in a repeatable way.
- Success is highly dependent on the particular data or subject matter (for example when the content of the tasks includes highly variable parameters, such as when booking flights).

Based on these criteria, it was concluded that the following are examples of products that could in principle be tested reliably:

- Fire extinguisher.
- Butterfly identification website.
- Pregnancy scanning machine.
- House paintbrush.
- Camera: for common goals only.

- A simple ecommerce shopping website (e.g., bookshop/rental car booking).
- Sewing machine (although quality criteria are hard to define).
- Oven - for the simpler task of reheating food only.
- Mobile phone - only for simple and well-defined goals.

The following products could not be tested reliably:

- Complex ecommerce website (e.g., purchasing a laptop, airline booking), excluding the less complex parts (e.g., basket and checkout) which do not include selection processes.
- Oven – for the more complex tasks of cooking food from ingredients.
- Microsoft Word – the content variability is too large and the quality of the main success criteria are not easily measurable (although it might be possible to test it for specific purposes).
- Car – too many goals and success measures (although car entertainment and navigation devices could be tested).

Note that the products that could not be tested reliably in their entirety, could be tested for specific goals e.g. the checkout part of a complex ecommerce site, or the starting of the car, or the creation of a new blank document in Microsoft Word, but in their entirety the products cannot be reliably tested.

In complex products like a flight booking web site, the usability is difficult to measure because it depends on the precise goal (for example there may be trade-offs between price and convenience), the particular range of potential alternative flights available to satisfy a particular query, and the features and ease of operation of the web site to support the particular goal. Although it is possible to compare aspects of the usability of different airline booking sites, there may be a temptation to select unrealistic or unrepresentative simple benchmark tasks, whereas in reality different users may have different complex goals. One could give examples of specific tasks that are easy or difficult to achieve on different sites, but it would be very difficult to obtain reliable and consistent measures.

In the first instance, the scope of 20282-2 will be for consumer products that do not have complex goals, but additional parts of 20282 could be produced for other types of products and systems that do not have complex goals, such as some types of web sites.

6 Assessing Goal Achievement

6.1 Defining Clear Success Criteria

Even for consumer products that are not "complex", one challenge is to find clear criteria for goal achievement. The criteria for what constitute successful achievement of a goal should match as closely as possible the criteria that would be applied by a typical user, rather than be limited by any constraints of the technology.

For some consumer products, there may only be one acceptable result (e.g., setting an alarm correctly). For others there may be a range of results that are acceptable to the user (for example a range of temperatures on an oven, or dryness of clothes from a tumble dryer).

The experts in the ISO group discussed the method that they would recommend using when evaluating the usability of an oven, which produced a surprising range of approaches to assessing goal achievement:

a) As the ultimate goal is to cook food correctly, some people suggested that the success measure should be that the food is cooked acceptably (e.g. a cake). Some thought that the quality of cooking should be judged by an expert, some that it should be judged by each user, and some that the method would be unreliable without an objective test.
b) Others thought that the goal should only be to set the oven to the intended temperature, but as the controls might be inaccurate, the actual temperature of the oven should be independently measured using a thermometer.
c) Another group thought that only the correct setting of the oven controls needed to be assessed (i.e. the scope of the testing should be the user interface only, rather than the whole product)

These issues with defining clear success criteria raise a much wider issue of scope: should the scope of the standard just be ease of use of the interface (called "ease of operation" in the current 20282-2), or should it be achievement of the user's goal (as in the ISO 9241-11 definition of usability)? For example, is an oven usable if it cooks at 200°C when set to 180°C? Or is an oven usable if it cooks at the correct temperature but the cakes are burnt at the end of cooking?

6.2 Who is Best Qualified to Judge Success?

In cases where the success criteria depend on some level of judgment, a new question arises: who should judge whether a goal has been achieved? For example, when cooking cakes, should it be an experienced cook or each user?

a) If expert judgment is used to assess goal achievement, would two experts make the same judgment?
b) If individual users are to judge, this introduces a new variable into the measures, and factors such as cultural variation could cause different tests to be incomparable. On the other hand, if experts are to judge, a standard measure of acceptability must be agreed on, and industry standard measures do not always exist, depending on the product being tested.

This issue of who should judge the results is compounded by potential variability in the skill or experience of the user. For example, some irons or sewing machines may make it difficult for less experienced users to produce acceptable results. In most cases the target audience will includes a mix of experienced and inexperienced users, and in order to be a realistic test, the variations in experience of these groups should be accounted for. Since these groups will have different expectations not only in usage, but also in quality of output, this further complicates the question of who should judge success.

6.3 Technical Tests and Usability Tests

An investigation into the testing methods used by the UK Consumers Association [5] revealed the following tests are performed:

- Oven: technical tests of heating accuracy, and expert assessment of ease of use and quality of cooked food.
- Camera: technical tests of image quality and ease of operation of the interface.
- Iron: expert review of speed of ironing and quality of results, and user review of ease of use.
- Sewing machine: expert and novice reviews of ease of use, and expert reviews of quality of results for experts and novices.

In these examples there are two independent sets of tests:

a) Technical or expert assessment of the capability of the product to produce acceptable results.
b) Expert and/or novice ratings of the ease of use.

These two separate tests should in principle produce results that are just as reliable as a single combined test, except when there is interaction between user operation and successful goal achievement. For example, some cameras may only achieve sharp telephoto pictures when held very steadily, while others (with an effective image stabilizer) will also achieve good results despite camera shake. In this case, measuring the ease of use of the controls independently of the technical quality of the photos would be insufficient to assess the ability to achieve the goal of a clear photo.

The potential inaccuracy in such cases could be circumvented either by performing a technical test to simulate the effect of camera shake, or by assessing the quality of photos produced during real use.

6.4 Approaches to Measuring Quality as Part of Usability

To resolve these problems, the ISO group is currently considering inclusion of the following procedure in the standard:

a) If a goal has only one outcome (for example, turning a product on) or has easily identified outcomes (for example, being woken at a particular time), achievement of the intended outcome shall be used as the basis for goal achievement.

   For example, the goal would not be met if a user believes that they have successfully set an alarm to ring at a particular time, but because they omitted the final step of pressing the alarm-on button, have failed to achieve the goal.

b) If the adequacy of the intended outcome of using the product can depend on how the product is used or operated (for example, if the quality of a picture is influenced by the steadiness with which a camera is held), the following tests for goal achievement should be considered in order of preference:

1. Carry out technical tests of the adequacy of goal achievement, if this is cost-effective and there are standards or other published criteria for what constitutes an acceptable level of quality in the resulting outputs.
2. Carry out expert assessment to define the quality criteria for adequate goal achievement, if the results would be reliable and cost-effective. These criteria should be based on standards or other published criteria for technical results that fall within an acceptable range.
3. Ask users to assess the acceptability of the quality of the outcome for defined purposes, if this is easy for a user to judge consistently by inspection.

Each of these methods has advantages and disadvantages, in particular, different users may use different criteria, and when using technical criteria these can be unwieldy and costly, and may not cover the same range as the criteria a user or expert would use.

For example, the quality of a photo could in principle be assessed by technical tests of the sharpness and color rendering, or by assessment by an expert, or by asking the user to assess the adequacy for a specific purpose (for example producing an A4 print). The quality of ironing could in principle be assessed by expert assessment, or by asking the user to assess the adequacy of the smoothness of particular clothes for a specific purpose.

The disadvantage of providing three options for the methods that can be used is that independent users of the standard might use different test methods. However, whenever comparisons are needed, the same method could be selected.

6.5 Big or Little Usability?

One objective of revising the standard was to broaden the scope from "ease of operation" (ease of use of the interface) to usability (effectiveness, efficiency and satisfaction when achieving goals). These could be regarded as "big" and "little" usability, referring to the wide scope versus the narrow scope.

The analysis above indicates that in some cases, in order to assess effectiveness reliably (as is required to cover "big" usability) complex technical tests may be required. These technical tests may be part of a broader test program (for example by a consumer organization), but could significantly increase the costs if this is not the case, and usability was being tested in isolation. In such situations, it may be desirable to use the standard to limit the scope to "small" usability, i.e. to just test the usability of the user interface, determining whether the controls can be set correctly regardless of the technical quality of the results.

While this type of testing is consistent with the popular interpretation of usability as ease of use, it does mean that even if a product is in this sense usable, it might also be useless, as it may not be able to technically perform the functions it was designed for. One objective of the definition of usability in ISO 9241-11 was to promote the practical and business relevance of usability interpreted as the quality of the product from a user perspective [1].

7 Conclusions

The current ISO 20282-2 standard contains a more rigorously specified procedure for obtaining reliable measures of usability than is available elsewhere, and is even more specific in the revised standard. The ISO 20282 series is intended to give acquiring organizations the opportunity ask for evidence of the usability of walk-up-and-use products and products for consumer use, and to give manufacturers the opportunity to publicize the results of usability test results. The method could also be adopted for use in certification schemes.

Some aspects of the method are still being developed, and academic and research organizations are invited to contact the ISO group (via the authors) to arrange trial use of the method, particularly to assess the consistency of results when the method is carried out by different test organizations.

Commercial organizations are invited to contact the ISO group (via the authors) to discuss the potential benefits of adopting the method.

The content of the standard is expected to be finalized and the revised standard published by ISO by 2013.

Acknowledgements. Thanks to Timo Jokela for helpful clarifications, and to the other contributors to the standard that have recently included: Tomas Berns, Susan Harker, Ran Linghua, Zhengjie Liu, Hans Persson and Lonneke Spinhof.

References

1. Bevan, N.: Quality in use: meeting user needs for quality. Journal of Systems and Software 49(1), 89–96 (1999)
2. Bevan, N., Schoeffel, R.: A proposed standard for consumer product usability. In: Proceedings of 1st International Conference on Universal Access in Human Computer Interaction (UAHCI), New Orleans (2001)
3. Bevan, N., Claridge, N., Maguire, M., Athousaki, M.: Specifying and evaluating usability requirements using the Common Industry Format: Four case studies. In: Proceedings of IFIP 17th World Computer Congress, Montreal, Canada, August 25-30, pp. 133–148. Kluwer Academic Publishers, Dordrecht (2002)
4. Carswell, C.M., Lio, C., McNally, J.: How Knowledgeable are Salespeople about the Usability of Their Merchandise? In: Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, v. 48, pp. 990–994 (2004)
5. Consumers' Association, http://www.which.co.uk
6. Ellison, C.: ViaMichelin X-930. PC Magazine. 2091601,00.asp (2007), http://www.pcmag.com/article2/0,2817
7. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11 Guidance on usability (1998)
8. ISO TC159/SC4/WG11: Ease of operation of everyday products (2011), http://www.iso.org/iso/iso_technical_committee.html?commid=53372
9. ISO/IEC 25062: Software Engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Common Industry Format (CIF) for Usability Test Reports (2006)
10. ISO/TS 20282-2: Ease of operation of everyday products – Part 2: Test method for walk-up-and-use products (2006)
11. Jokela, T.: When Good Things Happen to Bad Products: Where are the Benefits of Usability in the Consumer Appliance Market? ACM interactions XI.6, 28–35 (2004)
12. Overton, D.: 'No Fault Found' returns cost the mobile industry $4.5 billion per year, http://www.wdsglobal.com/news/whitepapers/20060717/20060717.asp
13. Sauro, J.: Confidence Interval Calculator for a Completion Rate (2010), http://www.measuringusability.com/wald.htm
14. Withers, P.: Fonehouse revamps store look and brand. Mobile News Online (2011), http://www.mobilenewscwp.co.uk/2011/02/fonehouse-revamps-store-portfolio-and-brand

Effects of Physical Display Size on GUI Designers' Perception and Implementation of Usability Guidelines

Selvihan Nazlı Kaptan[1] and Mehmet Göktürk[2]

[1] Bahçeşehir University, Çırağan Caddesi, 34353 Beşiktaş, İstanbul, Turkey
[2] Gebze Institute of Technology, Gebze, 41400, Kocaeli, Turkey
syavuzer@bahcesehir.edu.tr, gokturk@bilmuh.gyte.edu.tr

Abstract. Recent advances in technology led to common use of large screen in daily use. Applications of large displays are categorized in three groups; private, semi-public and public applications. Private and semi-public displays allow visualization and manipulation of large amounts of data at once. The increase in display size leads to an increase in data amount, which in result causes users to adopt a complex way of use. These changes arise the need for research in usability, interaction and user performance aspects of large displays. This study evaluates the effect of physical display size and resolution on GUI designers' perception and implementation of usability guidelines. Results suggest that subjects perform better on large displays for visual search and comparison tasks such as checking alignment, visual format similarities, etc. Reading related evaluations, on the other hand, are not supported by the larger displays.

Keywords: large displays, physical screen size, user performance, usability.

1 Introduction

Over the last decades, we have witnessed the establishment of numerous HCI related standards. In today's technology where different types of systems and users are present, users' expectations have increased while consistency and usability in general have become significantly important. While display size have been studied and evaluated in terms of individual task performance, the effect of display size on formation of usable interface design should also be discussed since it is another factor that affects the user performance. This study addresses the relation between the physical screen size and interface designers' ability to implement and evaluate user interface (UI) standards. If a limit can be defined on the screen size which enables designers to create more usable interfaces, one may guarantee more successful products from usability point of view. In the user experiments, we focused on two basic tasks that users perform on computers almost every day: Reading&comprehension and visual search&comparison tasks.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 128–137, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Background

The amount of data increases with the display size which, as a result, causes the users to adopt a complex way of use. However, too little is known on the differences between the users' behaviors and the physical and mental methods they adopt when they use such large displays and standard screen sizes. In studies [1, 2, 3, 4] on relation between display size and user performance, users have performed better on larger displays in spatial tasks, cognitive map formation and memory tasks. Bezerianos&Balakrishnan[7] and Huang&Mynatt [8] focused on determining the users' preference on screen size in their study and concluded that users prefer larger display sizes.

Independent from media, reading process involves not only resolution of information but also decoding the location of information in 2D space [5]. While reading on the computer screen, the reading process may be interrupted by navigational needs (such as paging, panning, etc). Field of view (FOV) should also be taken into consideration while evaluating screen size related factors. FOV plays a significant role on reading and comprehension ability. In [6], Kawashima suggests that the number of characters presented to user at once directly affects the comprehension performance of users. Virtual environment experiments shows there exists an upper limit on the FOV to ensure user success. Experiments with different FOV setups (10º, 20º, 40º, 80º and 120º) have shown users perform best when the FOV is 80º.

Visual search is another common task performed on computer screens. It involves focusing attention to different regions of view field and searching the focus area for a target object. Assuming the number of elements presented on the screen is constant, using larger displays will cause information spread over a larger area, which will cause more eye movements in pre-attentive phase and increased amount of data sent to brain. This will result in higher cognitive load and longer response times.

In the rest of this paper, reading&comprehension and visual search&comparison tasks will be used to evaluate the effect of physical display size on designers' performance on interface conformance tests.

3 User Experiment

In this study, we present the results of user experiments designed to reveal effects of display size and resolution factors on designers' ability to implement and evaluate user interfaces according to predefined design guidelines. During the experiments, designers were asked whether GUI forms that were given conformed to some predefined standard and the results were evaluated based on the response tie, correct response rate, mouse movement distance and participant behaviors.

User experiments had 3 different independent variables:

- Physical screen size
  - 17”
  - 22”
  - 40”
- Screen resolution
  - 800x600
  - 1024x768
  - 1280x1024
- Task Type:
  - Reading&Comprehension
  - Visual Search&Comparison

As physical screen size and resolution are varied, hypotheses in user experiments can be summarized as follows:

- ***H1:*** Participants will recall more information on larger displays.
- ***H2:*** Reading& comprehension will not be affected by physical screen size prvided that FOV will not be too high or too low. However, when the resolution is low, users’ reading performance will be lower on large displays.
- ***H3:*** On large displays, participants will perform spatial processing faster, and thus response times to related questions will be lower. However, correct responses will not change significantly.

3.1 Experiment Setup

In the experiments, 3 (three) different screens have been used. 1 17” Acer AL1716, 1 22” LG W2242S-PF LCD and 1 40” Alba LCD TV. On each display three different levels of resolution have been experimented.

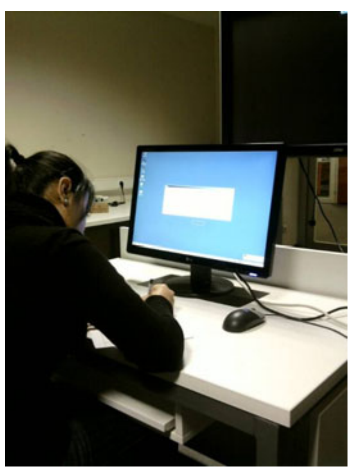


Fig. 1. Experiment setup

The experiment setup consisted of a display, an office table (except for the 40"), a mouse and a user chair. Since no keyboard input is required, no keyboard is placed on the experiment table for user comfort (Fig. 1).

During switching between screens, unused screens have been removed from the place, and the new screen has been placed at a marked distance from the user chair to preserve a constant FOV. In all experiments, displays were placed at proportional distances to preserve a FOV of approximately 50^o. Since previous studies have shown that user performance is higher between 45^o-60^o. While users' FOV was kept within this range, no method to fix FOV has been employed and users were left free to imitate a real working environment and to observe their responses to displays and resolutions.

3.2 Participants

10 participants (3 female, 7 male) have performed 3 experiments each to a total of 30 experiments. All the subjects have been chosen among graduate students who have worked on an interface design in the last 6 months. The participants consisted of 7 computer engineers, 2 software engineers and 1 electrical engineer.

3.3 Method

Each participant is first given a briefing on how the experiment will proceed. The procedure they should follow is explained in detail. Before the experiments, each participant has been informed that they will be given different-sized displays with different resolutions and that they will be requested to perform certain tasks on these display setups. They are asked to respond to these requests as quickly and correctly as possible with no given time constraints.

In each experiment, users are given a welcome screen with 5 design guidelines presented on it. Participants are given time to read and understand these guidelines and when they are ready they click a 'Start' button to start the experiment. In each question, a simple form is given and participants are asked whether the given form conforms to only one specific guideline given before. Each guideline involves either a reading&comprehension task or a visual search&comparison task. For example; users are asked to determine if there are any misspelled words in field labels, to determine whether all labels are aligned to same direction. Or to determine whether a given action button is present in the given user interface. On each screen, users are required to evaluate the given screen according to a single guideline only, and these guideline texts are given on each task screen so the users do not have to memorize them. Users are given two action buttons; if the interface conforms to the given guideline, the user clicks YES, otherwise user presses NO.

3.4 Collected Data

During the experiments participants and their screens have been recorded. Prior to experiments, participants have been informed that they are been recorded (Fig. 2).

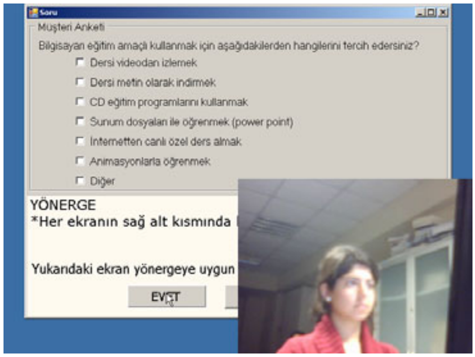


Fig. 2. User screen and camera capture

In addition to user recording, each participant's mouse movements, mouse total mouse displacement during the entire experiment, responses to questions and response times have also been recorded (Fig. 3).

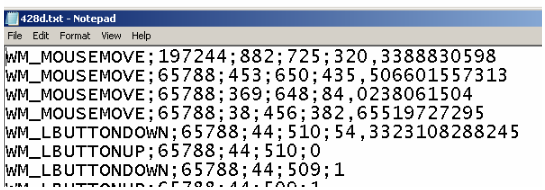


Fig. 3. Mouse movements record sample

Upon completion of computer-based experiment, participants are given a shot survey and they were asked to remember the guidelines they evaluated during the experiment. In addition, users scored and commented the experiment setup (display size and resolution) based on its convenience in daily work.

4 Results

Data collected during experiments have been analyzed according to task types, display sizes and resolution settings.

As expected, Figure 4 shows reading&comprehension response times do not significantly differ on 17" and 40" monitors. 22" monitor experiments resulted in decreased response times. This is a result that requires further analysis and reproducibility tests.

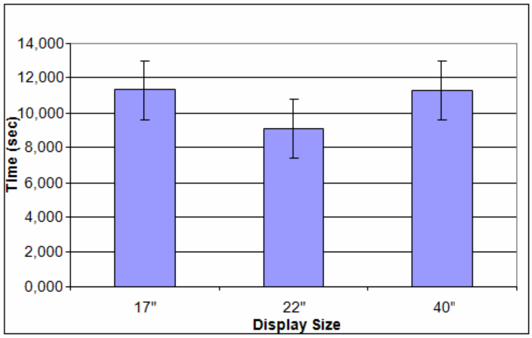


Fig. 4. Reading&Comprehension tasks average response times on different display sizes

On the other hand, Figure 5 shows that as display size increases, subjects became faster in cognitive processing and thus response times have reduced.

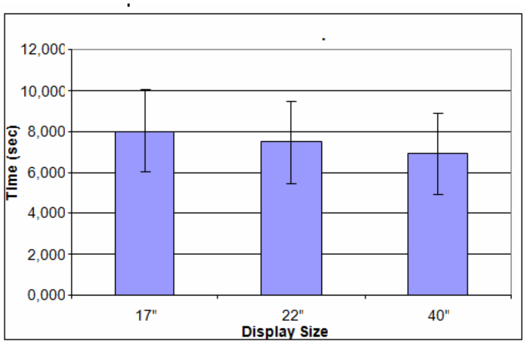


Fig. 5. Visual search&comparison average response times on different display sizes

In Figure 6, it is shown that different screen sizes and resolution levels had no significant effect on overall task success percentages.

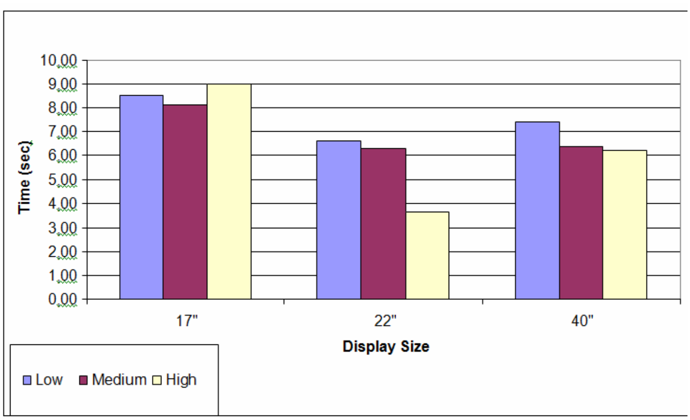


Fig. 6. Search&Comparison task response times on different display sizes and resolutions

With varying resolution levels, lower response times have been observed for search&comparison tasks with higher resolution settings on medium and large displays (Figure 7).

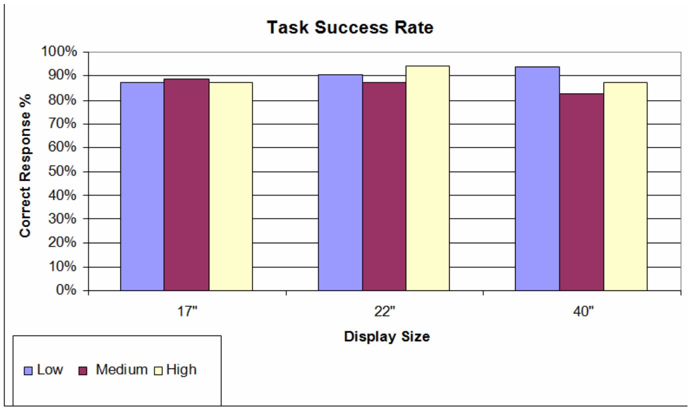


Fig. 7. Task success rates on different display sizes and resolutions

Since performance is both related to the correctness and the timing of the response, correct response per unit of time is used to evaluate the experiment results with a combined performance criteria. Simply, number of correct responses per unit of time is calculated as total correct responses divided by total response time. As can be seen from the Figure 8, resolution and display size has no direct effect on correct response per unit of time, with 22" high resolution, subject have shown the highest performance. However, according to subject surveys, this result may be related to the fact that most participants use 21"-22" monitors on a daily basis, so response times may benefit from subject familiarity.

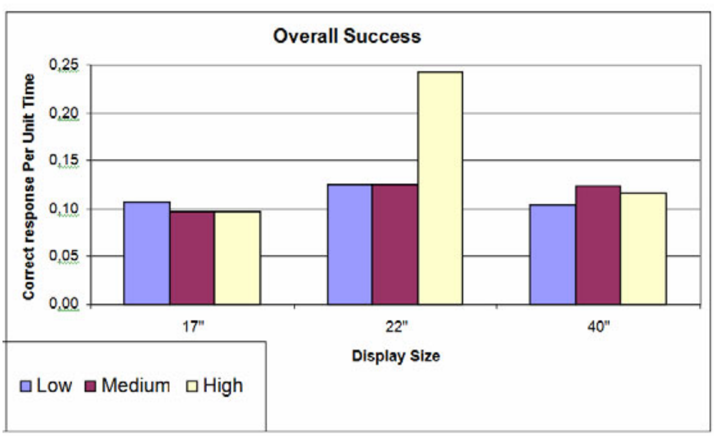


Fig. 8. Overall success rate on different display sizes and resolutions

4.1 Subject Comments and General Observations

Although subjects were not fixed to achieve constant field of view, in the experiment recordings no extreme movements to adjust sitting positions has been observed. This is important since large displays may be criticized for physical discomfort. While participants have not moved towards or away from the screen, they used mouse cursor as an aid to track their focus point. Although the experiments involved use of mouse for only clicking command buttons, subjects on large displays used mouse cursor for other purposes such as to check the alignment of GUI components and to follow text while they are reading.

After the experiments, participants were given a survey to score the experiment setup on a 0-4 scale where 4 is "strongly agree" and 0 is "strongly disagree". The results are given in Table 1 and Table 2.

Table 1. User preferences based on display size

Display Size	Would you use this display for daily work?	Would you use this display for interface design purposes (web/desktop interfaces)?
17"	4,00	3,67
22"	3,95	4,09
40"	2,10	2,40

Table 1 shows participants prefer 17" monitors for daily tasks, while 22" monitors are slightly more preferred for design purposes. 40" large display, on the other hand, has not been considered appropriate for either purpose. Similarly, in Table 2, medium and high resolution is more accepted by the users when compared to low resolution.

Table 2. User preferences based on resolution

Resolution	Would you use this resolution level for daily work?	Would you use this reslotion level for interface design purposes (web/desktop)?
800x600	3,00	3,06
1024x768	3,50	3,67
1280x1024	3,45	3,35

In these surveys, subjects have noted that 40" large display gives the feeling of watching a job done rather than actually doing it, so they have found it difficult to concentrate while working on a large display. Although this contradicts the idea of presence large displays claim to provide, these comments can be attributed to the nature of the given task –evaluating a GUI form- which can not benefit from presence.

5 Discussion

Users who have long preferred standard sized monitors, now prefer the largest displays they can afford and fit into their workspace. Under these circumstances, usability measures and interaction methods should be reevaluated. The fact that research have shown that screen size has direct effect on users performance brings up the problem of determining which screen size is more efficient on certain types of applications. In applications such as e-learning, information retrieval, crisis management and interface design, it may be possible to relate the screen size and user efficiency in order to match tasks with appropriate screen sizes.

While numerous studies have focused on 3D navigation and reading tasks on large displays, less has been done to observe the effects of physical display size on other frequently used skills. Goal behind this study is to observe reading&comprehension and 2D search&comparison processes under different display size and resolution settings. In some cases, results have confirmed the reporting of previous studies. Still, the experiments show that the amount of information subjects remember after testing on a large display is less when compared to subjects who tested on smaller displays. There are experiments that show users of large displays remember more information in 3D tasks. This indicates that cognitive processes involved with different tasks are affected by the screen size in different ways.

References

1. Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., Tan, D.: The large-display user experience", Computer Graphics and Applications. IEEE Computer Graphics and Applications 25(4), 44–51 (2005)
2. Czerwinski, M., Tan, D.S., And Robertson, G.G.: Women Take A Wider View. In: Proceedings Of The Chi 2002 Conference On Human Factors In Computing Systems, pp. 195–202 (2002)
3. Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J.A., Vewatti, T., And Chiselko, G.: Using a large projection screen as an alternative to head-mounted displays for virtual environments. In: Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems, pp. 478–485 (2000)
4. Tan, D.S., Gergle, D., Scupelli, P., Pausch, R.: Physically large displays improve performance on spatial tasks. ACM Trans. Comput. 13(1), 71–99 (2006)
5. Wästlund, E., Norlander, T., Archer, T.: The effect of page layout on mental workload: A dual-task experiment. Comput. Hum. Behav. 24(3), 1229–1245 (2008)
6. Kawashima, H., Oda, K., Kobayashi, A., Suzuki, M.: Does limited window size affect text comprehension? In: Vision 2005 - Proceedings of the International Congress held between April 4 and 7, 2005 in London, UK, September 2005. International Congress Series, vol. 1282, pp. 622–626 (2005), ISSN 0531-5131
7. Bezerianos, A., Balakrishnan, R.: View and space management on large displays Computer Graphics and Applications. IEEE 25(4), 34–43 (2005)
8. Huang, E.M., Mynatt, E.D., Russell, D.M., Sue, A.E.: Sue, Secrets to Success and Fatal Flaws: The Design of Large-Display Groupware. IEEE Computer Graphics and Applications 26(1), 37–45 (2006)

9. Bystrom, K.E., Barfield, W., And Hendrix, C.: A conceptual model of the sense of presence in virtual environments. Presence: Teleoperators and Virtual Environments 8(2), 241–244 (1999)
10. Lin, C.J., Hsieh, Y.H., Huang, C.W.: Task performance and visual discomfort while working with a large visual display unit. In: Asia Display 2007 International Conference (AD 2007), Shanghai, PEOPLES R CHINA, pp. 953–958 (2007)
11. Tyndiuk, F., Lespinet-Najib, V., Thomas, G., Schlick, C.: Impact of large displays on virtual reality task performance. In: Proceedings of the 3rd international Conference on Computer Graphics, Virtual Reality, Visualisation and interaction in Africa (AFRIGRAPH 2004), Stellenbosh, November 03 - 05, pp. 61–65. ACM, New York (2004)

Terminological Precision - A Key Factor in Product Usability and Safety

Barbara Inge Karsch and Gabriele Sauberer

International Network for Terminology - TermNet
Mooslackengasse 17, A-1190 Vienna, Austria
bikterminology@gmail.com

Abstract. Precise use of terminology must be a key component in the communication amongst the product team and with end users. It is not a simple goal to have standardized terms used throughout a lengthy development process in which many people with different expertise and at different locations are involved. Terminology tools and processes used by trained terminologists enable precise use of terminology throughout the product life cycle, in all content management systems and by all contributors to the content supply chain.

Keywords: Terminology, precision, usability, terminology management system, life sciences.

1 Introduction

The natural sciences underlying the concepts in the life science industry are precise: Biology, anatomy, physiology, etc. are based on systematic research and documentation, e.g. Carl Linnaeus's nomenclature for plants, the animal kingdom, and the periodic table of elements. These sciences need organization, as organization enables precision. When a pharmaceutical company is developing a new drug or when a medical device maker is designing a new diagnostic tool, precision is a crucial component. Without precision, drugs or devices might not achieve what they are intended to, might need to be recalled, or might even harm a patient.

Precision must also be a key criterion for communication, for instance, with a patient using a medical device. Accurate communication on a small user interface depends on the correct and consistent use of technical terminology. It is not a simple goal to guarantee the usage of standardized terms throughout the device development process, though: Many contributors are involved; development might take months, if not years; documents are prepared in different systems, by different stakeholders, and in different locations.

Well-defined naming and solid documentation practices, a trained stakeholder in charge of maintaining technical terminology and a terminology management system (TMS) accessible to all project participants are critical for consistent and user-centered terminology. They provide the organization that enables reliable delivery of seemingly simple, yet highly complex products.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 138–147, 2011.
© Springer-Verlag Berlin Heidelberg 2011

After two illustrations of problems caused by unmanaged terminology in a life science field, this paper presents how precise communication with end-users can be successfully supported through terminology management practices. It describes what terminology management is and how it correlates to the product development cycle.

2 Definitions

Precision is a term with many meanings. For the purpose of this paper, precision is defined as “the quality of being sharply defined by virtue of exact detail” [1].

The language used on the user interface (UI), in supporting documentation or in instructional material falls into the category of language for special purposes. LSP is the “language used in a subject field and characterized by the use of specific linguistic means of expression [which] always include subject-specific terminology and phraseology” [2]. Examples for subject fields in life science products are medical specialties (e.g. cardiology), but also computer science in cases when the terminology appears on or refers to a device UI.

The main vehicle for transferring knowledge from a device or pharmaceuticals manufacturer to the patient is terminology. Terminology is defined as “the collection of designations, i.e. terms, appellations and symbols, belonging to a language for special purposes” [2]. Examples are ‘sphygmomanometer, ‘Federal Drug Administration’ and ⚠, respectively.

As shown in the semiotic triangle (see in [3] and others) below, a designation, for example ‘sphygmomanometer’, represents a concept. In order to communicate effectively about concepts, humans agree on definitions (“A pressure gauge…” [4]).

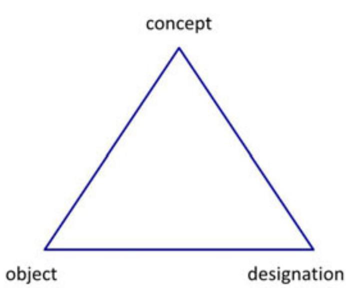


Fig. 1. Semiotic triangle

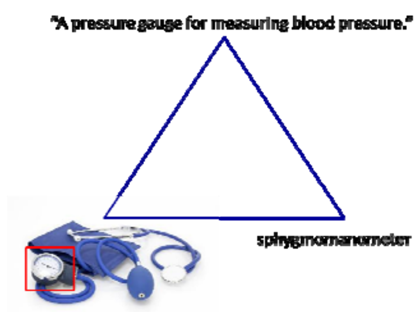


Fig. 2. Semiotic triangle with an example for a term, the corresponding object [5] and definition [4]

Terminologists use terminology principles and methods to document concepts and their corresponding designators in databases for reuse by others in written and spoken communication. The goals of managing terminology are, among others, to enable fast and safe use of a product, and to minimize additional research, misunderstandings and ultimately cost.

3 Why Terminology Management

While misunderstandings in every-day language can lead to amusing outcomes at best, misapprehensions in LSPs can result in production delays, in recalls and potentially in disasters. As the product moves through the different product life cycle stages, content is handed off to different people. R&D might have included writers and editors, for example, but even these content publishers are slightly more removed from the conceptual phase. They will hand off finished or semi-finished content to translators who are even further removed from the conceptual origin of a new device.

All of these "senders of communication" face the challenge of writing or speaking about product concepts clearly and consistently and in the language of the receiver, so that the message can be understood. Creating clear and consistent strings on a device with a user interface is a particularly daunting task, as space is limited. UX design expert, Everett McKay, not only stresses the importance of user-centered design, he says "[f]ocusing on effective communication is the single best user-centered design technique [6]".

Without a terminology strategy, problems are inevitable. Misnomers and inconsistencies are two of the most common problems. Incorrectly labeled functions, inadequately abbreviated buttons or simply poorly motivated terms are examples of misnamed concepts. Users may be slowed down in the use of the product, may need to research before using the product, may not use the product or may use it incorrectly. Specifically the latter presents a high risk for manufacturers of life science products.

According to a recent survey conducted by the International Network for Terminology, TermNet [7], inconsistent terminology is the number one factor impeding writers, editors, translators, etc. in the US information and communication technology (ICT) sector. 83% of the survey respondents said that inconsistencies between documents or within documents are consuming their time. Three participants even said that they have experienced product recall due to terminology issues. The following are examples of faulty terminology leading to incorrect study results and to a recall.

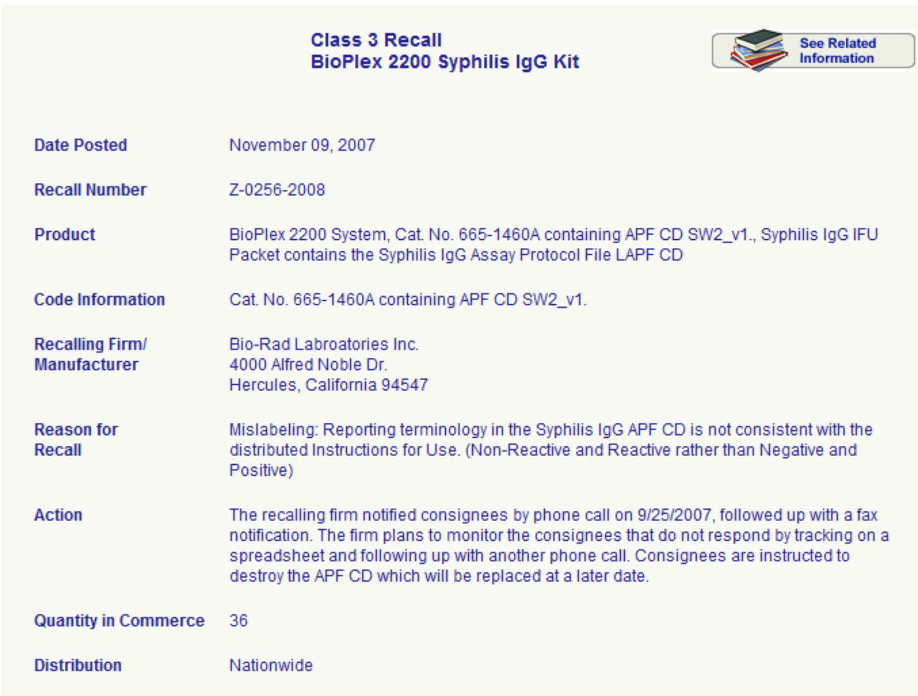
3.1 Misunderstandings

The Redbook guidance document by the Food and Drug Administration [8] recognizes the importance of defined and standardized terminology. The section "Inconsistency in Applying Diagnostic Terminology" lists the following real example:

"A study was submitted in which tissues from about one-third of the animals were evaluated by the study pathologist and the remainder were evaluated by a consulting pathologist. The diagnostic terminology was not consistent between pathologists and no attempt was made to explain the inconsistencies in the study report. Although the data appeared to show treatment-related effects, these were subsequently attributed to the way different categories of lesions were summarized." [8]

3.2 Recall Case

The FDA's recall database for medical devices contained the following instance of a recall due to a terminology error in January 2011. As Figure 1 shows, "reporting terminology in the Syphilis IgG APF CD is not consistent with the distributed Instructions for Use. (Non-Reactive and Reactive rather than Negative and Positive)"[9].


Class 3 Recall
BioPlex 2200 Syphilis IgG Kit

See Related Information

Date Posted	November 09, 2007
Recall Number	Z-0256-2008
Product	BioPlex 2200 System, Cat. No. 665-1460A containing APF CD SW2_v1., Syphilis IgG IFU Packet contains the Syphilis IgG Assay Protocol File LAPF CD
Code Information	Cat. No. 665-1460A containing APF CD SW2_v1.
Recalling Firm/ Manufacturer	Bio-Rad Labroatories Inc. 4000 Alfred Noble Dr. Hercules, California 94547
Reason for Recall	Mislabeling: Reporting terminology in the Syphilis IgG APF CD is not consistent with the distributed Instructions for Use. (Non-Reactive and Reactive rather than Negative and Positive)
Action	The recalling firm notified consignees by phone call on 9/25/2007, followed up with a fax notification. The firm plans to monitor the consignees that do not respond by tracking on a spreadsheet and following up with another phone call. Consignees are instructed to destroy the APF CD which will be replaced at a later date.
Quantity in Commerce	36
Distribution	Nationwide

Fig. 3. Entry in the Medical Device Recall Database of the US Food and Drug Administration [9]

In this case, the terminological inconsistency does not even appear to be major, after all the faulty terminology is semantically very close to the intended terms. It is easy to imagine more grave errors. At a minimum, recalls cost time and money. They could cost reputation, licenses, and, if not carried out successfully, ultimately lives.

4 Terminology Management

Terminology management is the systematic research, documentation and reuse of concepts and their terms [3]. When new concepts are first developed and described, a terminologist helps the subject matter experts, e.g. a team of cardiologists, coin designations and documents information about it in a terminology management system (TMS). Anyone, from the research staff to the writer to the translator, can then access the centralized database and use correct terminology consistently. The following is a brief overview over skills, tools and processes necessary for successful research, documentation and reuse of standardized terminology.

Skills. Terminologists fall into two categories: The subject matter expert who has the technical knowledge of the subject area as well as enough awareness of terminology methods and principles; and the trained terminologist who has the research skills to quickly get into a subject matter [10]. There are two-year master's degree programs as well as certification programs (e.g. ECQA Certified Terminology Manager). Some of the most important terminology (vs. subject matter) skills are:

- Solid understanding of terminological and terminographical working methods
- Linguistic skills in one or more languages
- Tools and computer skills
- Problem solving and research abilities
- Communication, networking, and social competencies

Companies, such Scania, VW or Siemens, organizations, such as the European Patent Office or the European Commission, and national language centers in Finland, Sweden or Ireland have a terminologist or small teams of terminologists who support the content supply chain with expertise, for example, in term formation and terminology data management. To enable data interoperability, they are also familiar with terminology management standards set forth, for instance, in ISO documents created by Technical Committee 37.

Tools. Terminology management systems (TMS) are software applications that run on top of a relational database. One of the main criteria for a TMS is that it is concept-centered. That means that a concept, indicated by its unique ID and a definition, is the main entry. Attached to it are the designators that stand for the concept. In our example that would be the terms "sphygmomanometer" or "blood pressure meter," depending on the needs of the users and the agreed upon (=standardized) usage. It could also have the terms "Blutdruckmessgerät" for German or "tensiomètre" and "sphygmomanometer" for French attached to it in a multilingual setting.

The software interface is generally designed for the terminology expert who prepares data which is then used in content management systems, translation memory tools, machine translation systems, etc. While there are several commercial tools available, particularly software companies, such as SAP, IBM, Microsoft or Oracle, created proprietary solutions.

Processes. There are various ways to break down the terminology management process. The expert team responsible for the best practice guide of the German terminology association, *Deutscher Terminologie Tag e.V.*, lists the following four phases [11]: Terminology production, preparation and distribution, use, and maintenance.

- Terminology production is characterized by the identification of existing and the creation of new concepts. Terms that represent these concepts are collected, and new terms or names are created.

- During the terminology preparation phase, concepts and terms are being standardized and documented in a TMS along with metadata, such as definition, part of speech, context, or subject.
- As soon as the terminological entries in the database are stable, they can be released, and writers, editors, and translators among others access and use them.
- Some terminological entries might need to be changed because errors are detected or new information is available. This happens during a maintenance phase of an entry.

The main outcome of the process, established by applying the skills and in the terminology management system, is entries, such as the example in Figure 2 [12].

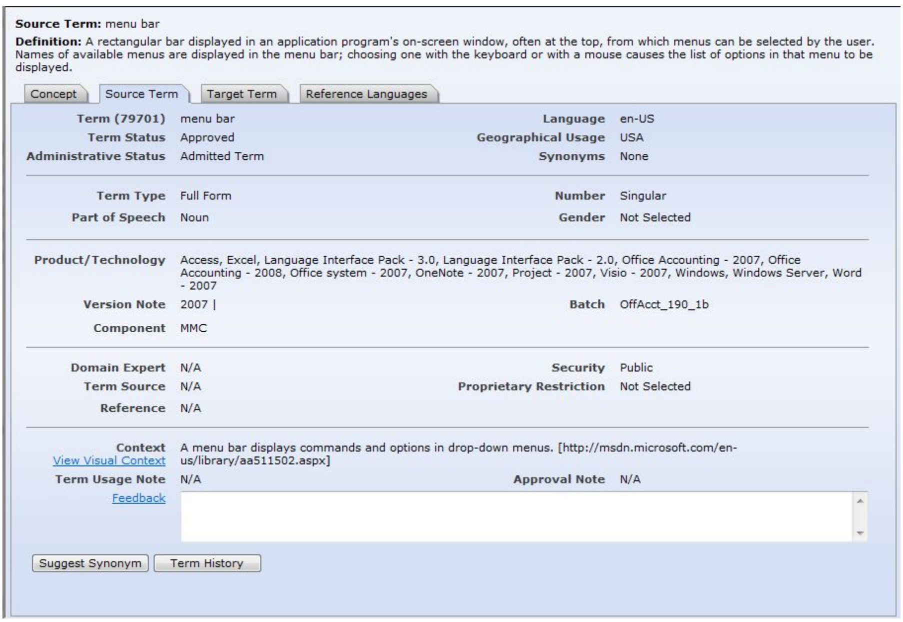


Fig. 4. Terminological entry for “menu bar” in a TMS [12]

5 Terminology and the Product Life Cycle

Depending on the industry and the device, product life cycles (PLC) take slightly different forms. This section describes how the product life cycle and the terminology life cycle are intertwined.

Figure 3 is a PLC model created by the FDA for medical devices. The phases Concept, Prototype, Preclinical, Clinical, Manufacturing, Marketing, Commercial Use, and Obsolescence are combined into Phases I, II and III below to facilitate discussion. During these phases, a variety of documents are created. The most common ones are: requirements documentation, design documentation, technical documentation, user documentation, and marketing documentation. Table 1 juxtaposes the approximate phases with the documents they produce.

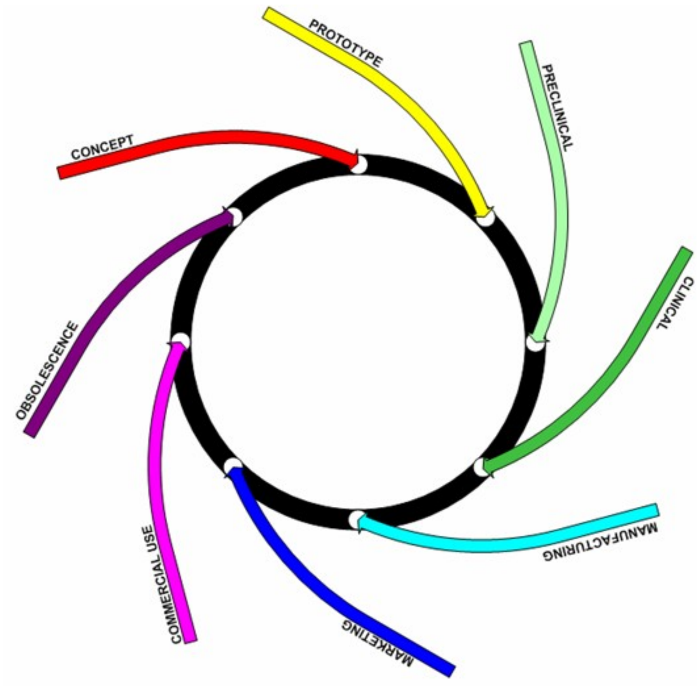


Fig. 5. Model for the total product life cycle by US Food and Drug Administration [13]

These documents contain technical terminology that is at different levels of stability during the PLC phases. Stability, for our purposes, is the quality of being free from change. Ideally, terms are stable from their inception; realistically, stability should be reached before a document is released either to the final consumer or the next step in the workflow (e.g. from content publishing to translation), so as to minimize costly changes. The next three paragraphs detail out the phases and their terminology cycle, followed by an overview in table format.

Table 1. Phases of the PLC and documentation created during these phases

Phase category	PLC Phase	Documentation
I	Concept	Requirements
	Prototype	Design
	Preclinical	Technical
II	Clinical	User
	Manufacturing	Marketing
	Marketing	
III	Commercial use	Marketing
	Obsolescence	

Phase I. Medical researchers, anthropologists, and other experts gather and analyze data. Terms and names designating these concepts might come up during patient interviews, field studies or literature reviews. They are used in drafts for requirements, design and technical documents with the understanding that they may still change. This is the phase were valuable conceptual expertise is present, but the product in all aspects, incl. terminology, is most in flux. Terminology stability is low.

Phase 2. Once problem analysis is concluded and development of a solution is well on its way, naming gets firmer. Specific attention must be placed on terminological choices on small user interfaces. Consistent and natural language is critical to usability. This is the phase when user and marketing material is written: Terms, names and labels must be locked in and documented in the TMS for everyone to use consistently. Now is also the time to research and document terms in other languages to prepare for translation. Pre-set terminology is distributed to all translators who use it during the translation of the product and all accompanying material. Questions regarding additional terms are answered and documented in the database.

Phase 3. During commercial use and obsolescence, tests for the next version of a product or a completely new product will start. Terminological problems or insufficiencies might emerge and should be noted in the terminology management system, so that better terms and names may be used in the future. Terms in the existing product are obviously not changed and stability is therefore high.

Table 2. PLC, the terminology process and stability

Phase	PLC Phase	Documentation	Terminology process	Terminology stability (ideal)
I	Concept	Requirements	Production	Low
	Prototype	Design	Production	Low
	Preclinical	Technical	Documentation, use and feedback	Medium
II	Clinical	User Marketing	Documentations, use and feedback	Medium
	Manufacturing		Use	High
	Marketing		Use	High
III	Commercial use	Marketing	Use and maintenance	High
	Obsolescence		Use and maintenance	High

6 Conclusion

Terminology errors can spoil lab data or lead to recalls. Even when they are caught before release of a product, they cause costly disruption of the content flow. And even if company terminology is locked down and documented, official FDA terminology might undergo changes, which might come during any phase of the PLC, worst of all just before release.

A terminologist may be a full-time or part-time role on the team. It may be a subject matter expert with terminology skills or a terminologist with subject matter expertise. But a database must be kept up-to-date to get the most return on investment. Depending on the length of the product life cycle, the number of products and languages managed in the database and the number of users it supports, coordination can be done by one person or a team.

Concept-based terminology management systems allow database users to select the correct term for the target audience of their message. That may be "sphygmomanometer" when experts are communicating about the gauge only; it maybe "blood pressure meter" when the content of a blood pressure kit is described in the user documentation; or it might be "the cuff" when a nurse refers to the measuring device to a patient. Precision is user-dependent.

When precision is not present, errors will occur, if not in the source language, very likely during the translation process. A translator relies on the source material to be precise and error-free. If the text reads "cuff" when the entire blood pressure kit is meant and it is not clear from the context, the translation will be faulty. It is of utmost importance that terms and names be tracked in a database from the concept phase on and used by everyone correctly and consistently.

The above is an oversimplified process flow that in reality is impacted by tight deadlines, unexpected changes in the development cycle, and the number of stakeholders. Appropriate workflows, supported by trained terminologists and well-designed terminology management systems speed up the development process, reduce documentation and translation cost, and last, but not least, lead to precise terminology. Precise terminology as part of a products user-centered communication results in safer and more usable products and devices.

Acknowledgments. The authors gratefully acknowledge the feedback of Wilbur Pierce, Ph.D.

References

1. Elsevier Inc. (ed.): Saunders Comprehensive Veterinary Dictionary 3, The Free Dictionary by Farlex, http://medical-dictionary.thefreedictionary.com/precision
2. International Organization for Standardization: International Standard 1087-1, Terminology work - Vocabulary - Part 1: Theory and application, Geneva (2000)
3. Arntz, R., Picht, H., Mayer, F.: Einführung in die Terminologiearbeit. Studien zu Sprache und Technik, 4th edn., vol. 2. Georg Olms Verlag, Hildesheim, Zürich, New York (2002)
4. WordNet Search, http://wordnetweb.princeton.edu/perl/webwn?s=sphygmomanometer
5. Medical Device Safety Institute, What is a Medical Device?, http://www.medicaldevicesafety.org/c17/What-is-a-Medical-Device?htm
6. McKay, E.: UI is Communication. Presentation, http://www.uxdesignedge.com/wp-content/uploads/2010/08/UI-is-communication1.pdf
7. Karsch, B.I.: Terminology Survey Results. Vienna, TermNet (2011)

8. Food and Drug Administration: Toxicological Principles for the Safety Assessment of Food Ingredients. Redbook 2000: IV.B.3 Pathology Considerations in Toxicity Studies, http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/FoodIngredientsandPackaging/Redbook/ucm078318.htm#ivb3bii
9. Food and Drug Administration. Class 3 Recall BioPlex 2200 Syphilis IgG Kit, Medical & Radiation Emitting Device Recalls Database, http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRES/res.cfm?id=65277
10. Engel, G., Picht, H.: Der Terminologe - Beruf oder Funktion? In: Hoffmann, L., Kalverkämper, H., Wiegand, H.E. (eds.) Fachsprachen - Languages for Special Purposes, pp. 2237–2244. Walter de Gruyter, Berlin (1999)
11. Reineke, D. (ed.): Terminologiearbeit - Best Practices. Deutscher Terminologie-Tag e. V., Cologne (2010)
12. Karsch, B.I.: Anforderungen an Terminologen bei Microsoft. In: Best Practices in der Terminologiearbeit, Deutscher Terminologie Tag e.V., Heidelberg (2010)
13. Feigal, D.W.: Total Product Life Cycle, http://zone.ni.com/devzone/cda/tut/p/id/5711

ISO Standards for Standard Software: Accountability, Customer Expectations and Reality

Ulrich Kreichgauer

SAP AG, Dietmar-Hopp-Allee 16,
69190 Walldorf, Germany
Ulrich.Kreichgauer@SAP.com

Abstract. Using the usability ISO product standard 9241/110 as an example, the author will show that standard software manufacturers can lay the foundation for the requirements described in usability ISO standards. However, the final fulfillment of these standards comes through customization commitments from the customer buying and using the standard software.

Keywords: Usability, ISO 9241/110, standard software, accountability, customization.

1 Introduction

ISO 9241 / 110 describes "dialogue principles" for usable software. By following the principles "the ultimate beneficiary will be the user" [1], because the ISO standard "will lead to user interfaces (UIs) that are more usable, consistent and that enable greater productivity" [1]. The standard specifies seven ergonomic principles which apply to the dialogue design between people and information systems:

- Suitability for the task
- Suitability for learning
- Suitability for individualization
- Conformity with user expectations
- Self descriptiveness
- Controllability
- Error tolerance

The standard is used to define software requirements and to check software according to those principles.

2 Importance of the Standard

While an ISO standard is not the same as a law, it can be enforced using a law. For example, with the ISO 9241/110 this is done in the European Union the following way:

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 148–153, 2011.
© Springer-Verlag Berlin Heidelberg 2011

The generic "Council Directive 89/391/EEC" of June 12, 1989 [2] defines measures to encourage improvements in workers' safety and health at work. For computer based workplaces the more detailed "Council Directive 90/270/EEC" of May 29, 1990 [3] defines the minimum safety and health requirements for work with display screen equipment.

Each country in the European Union implemented laws, based on those directives, addressing the issues in a binding way. In Germany for example, the "German Health and Safety at Work Act (German: Arbeitsschutzgesetz)" [4] defines which requirements all workplaces must fulfill to ensure safe and healthy work and the "VDU work directive (German: Bildschirmarbeitsverordnung)" [5] defines requirements for computer based workplaces concerning hardware and software aspects.

In more detail, the VDU work directive states in the "appendix about requirements for computer based workplaces (Anhang über an Bildschirmarbeitsplätze zu stellende Anforderungen)":

"20. The principles of ergonomics have to be applied, specifically concerning the processing of information by humans. (20. Die Grundsätze der Ergonomie sind insbesondere auf die Verarbeitung von Informationen durch den Menschen anzuwenden.)" [5]

None of the directives link directly to an ISO standard, but experts usually refer to the more concrete ISO standards, which describe for example the principles of ergonomics. These are ISO 9241 (Ergonomics of human-system interaction) [6], ISO 9355 (Ergonomic requirements for the design of displays and control actuators) [7], ISO 14915 (Software ergonomics for multimedia user interfaces) [8] and various others. By fulfilling the detailed requirements of the relevant ISO standards, the high level requirements of the mandatory directives and laws will also be fulfilled.

3 Compliance

In countries with similar laws, companies have to prove that computer workplaces are compliant with the directives, i.e. that they are safe and healthy. In addition to hardware (monitors, desks, chairs etc.), software also plays an important part in the compliance and is enforced as described above. Under such laws workers can refuse to work at non-compliant workplaces. However, works councils, where they exist, usually enforce the required compliance for the workers in advance.

The directives and laws target the workplaces at the companies, not the hardware and software manufacturers. If we take the scenario where software is designed specifically for one company, compliance to the directives and laws can be assured, since during development the software is designed to fit explicitly to that company's and its users' needs. In contrast to this scenario, standard software is built using a "one size fits all" approach, with built-in adaption capabilities. Standard software must be configured, i.e. “customized” and “personalized” to fit the needs of the specific company, groups of users within the company, and even individual users.

By applying "customization" to standard software, the software is adapted to the company's needs (e.g. by adding information about the company's structures, typical processes, company specific requirements). Customization settings are valid within the whole company and therefore influence all users equally.

Further adaptations are made by applying "personalization", i.e. the software is adapted to the needs of single users (e.g. by choosing font sizes, by setting personalized page views, by choosing column orders in tables). Personalization settings may vary between users, to support each user in his/her individual needs.

Nevertheless, companies buying and using standard software would like to have a simple and easy way to guarantee compliance. They often ask for a kind of certificate or statement from the software manufacturer that “the software is ISO / directive / legally compliant” or that “all aspects of the computer workplace usability are built into the software.”

However, due to the standard software’s built-in adaption capabilities, a global statement or certificate about the ISO compliance cannot be granted by the software manufacturer alone. After customization and personalization is completed at the customer site, the standard software may look, feel, and function very differently for individual users of one company, and even more so for individual users of different companies. Therefore, to guarantee maximum compliance, both the software manufacturer and its customers have to work hand-in-hand.

This work is reflected in the ISO 9241/110 principles. Some principles are more dependent on the manufacturer's development work, e.g. error tolerance, suitability for individualization, self-descriptiveness. Others can usually only be made in the software by the final customization and personalization efforts at the customer site, e.g. suitability for the task, conformity with user expectations. Some are a mixture of both, e.g. controllability, suitability for learning.

4 The Responsibilities of the Standard Software Manufacturer

To fulfill the requirements described in the ISO 9241/110 [1] in standard software a software manufacturer must build the software in an adaptable way. Using a "user-centered design process" the needs and tasks of typical users ("personas") need to be researched. Researching various end-users / companies / locations usually demonstrates which tasks and user needs are:

- Typically valid for all companies / users / business processes / business steps. These are usually implemented in a standardized way into the software.
- Slightly different between companies / users / business processes / business steps. These are usually implemented so that the standard software can be customized and personalized via switches, to ensure “best fit”.
- Very unique for specific companies / users / business processes / business steps. These are usually implemented in a way that dummy fields, dummy screens, etc. can be activated at the customer site; or separate add-ons, developed as custom projects, are plugged into the software and ensure a fine-grained fit.

The basis for this approach is the proper understanding of the end-user's tasks and needs, leading to the correct implementation of the software. This includes the adaptation options which of course must be documented. By providing this stable and flexible approach to the user interface, the customer can then make the necessary adaptations to the software to fulfill the requirements of the ISO 9241/110 standard.

5 The Responsibilities of the Customer

Finally (and only!) at the customer site must the concrete tasks and user needs be known. An "on-site" analysis of those tasks and needs must be done to ensure a proper software fit. With these facts and findings, the customer can and must customize the software accordingly, for example by showing necessary or hiding unnecessary fields, field groups, screens or functions, by arranging field orders, or by providing customer specific search or filter options. They must also train the users, not only how to use the software to do their job but also how to adapt the software using personalization options for individual users.

Execution of the tasks and needs analysis is usually the job of usability professionals. Specialized usability consulting companies should not only have methodological knowledge, but also an understanding of the customization and personalization options of the used software [9].

After applying all of these required steps, the ISO compliance should be evaluated, either by using checklists provided by the ISO standards, or by using checklists like ISONORM 9241/110-S [10, 11].

6 Examples

Using ISO 9241/110 requirements in examples, you can see what the concrete responsibilities are for both the software manufacturer and the customer.

6.1 Example 1

Suitability for the task requirement 4.3.5 states: “The steps required by the dialogue should be appropriate to the completion of the task, i.e. necessary steps should be included and unnecessary steps should be avoided.” [1] Suitability for individualization requirement 4.9.9 states: “If appropriate, it should be possible for the users to add or rearrange dialogue elements or functionality specifically supporting their individual needs when carrying out tasks." [1]

Software manufacturer. Based on user-centered design research, for a specific task and persona it is specified which:

- Required fields are needed in all variants of a business process (implemented as "fixed UI elements").
- Additional fields, field groups or screens are needed by specific business processes - (implemented in a way that they can be shown or hidden via customization).
- Additional fields, field groups or screens might be needed by company policies (implemented in a way that there are dummy fields, field groups or screens which need to be labeled according to the task or predefined APIs for add-ons providing specific fields or functions).

Customer. Based on customer, end-user and task analysis, additional fields, field groups or screens might be shown or hidden from users via customization and personalization. In addition, end-user adaptations can be provided. These could be color schemes, font-sizes, column orders or specific functions.

6.2 Example 2

Controllability requirement 4.7.5 states: "If the volume of data relevant to a task is large, then the user should be able to control the data presented." [1]

Software manufacturer. Based on user-centered design research, for a specific task and persona it is specified how "large volume of data" is presented. For example, this can be done by defining views, appropriate controls, filters and search functions.

Customer. Based on customer end-user and task analysis, user and task-specific filters, specific search options, additional search fields, specific default views, individual column orders, etc. should be defined and activated via customization and personalization.

7 Effect of Customization and Personalization Efforts

Efficiency is one of the usability criteria of ISO 9241/11, defined as "resources expended in relation to the accuracy and completeness with which users achieve goals" [12]. And the customization and personalization efforts at the customer site are especially important for efficiency as proven in a usability study, which compared non-customized standard software with optimized/ customized standard software [13]. Users working with the non-customized software needed two to four times longer to fulfill a task compared to users of the customized software. The amount of keyboard and mouse interaction for data input and navigation was as much as two times higher in the non-customized software.

This study relied only on customization. No additional programming was done, no add-ons were used, and no personalization was made. The study mentioned that by applying these additional adaptations even more efficiency could be reached.

8 Collaboration Model

Software manufacturers must of course ensure a high level of usability for standard software, even without customization and personalization. Nonetheless, it is in their interest to make the benefits of these adaptations transparent to their customers and easy to implement. If they don’t, the software may not offer enough flexibility or too much flexibility, making the adaptation effort too high for customers.

On the other hand, customers must realize that standard software is built to fulfill many companies’ and users’ needs. If they do not invest enough time and money to customize the software and train their individual users how to personalize it, even standard software with the best built-in usability will not fulfill usability’s primary goals of “efficiency, effectiveness and user satisfaction”.

References

1. ISO 9241/110: Ergonomics of human-system interaction – Part 110: Dialogue principles (ISO 9241-110:2006). International Organization for Standardization / TC 159/SC 4 - Ergonomics of human-system interaction (2006)
2. Council Directive 89/391/EEC of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work (1989)
3. Council Directive 90/270/EEC of 29 May 1990 on the minimum safety and health requirements for work with display screen equipment (fifth individual Directive within the meaning of Article 16 (1) of Directive 89/391/EEC) (1990)
4. German Health and Safety at Work Act (German: Arbeitsschutzgesetz / Gesetz über die Durchführung von Maßnahmen des Arbeitsschutzes zur Verbesserung der Sicherheit und des Gesundheitsschutzes der Beschäftigten bei der Arbeit vom 7. August 1996, zuletzt geändert durch Verordnung vom 31. Oktober 2006) (1996)
5. Work Directive, V.D.U. (German: Bildschirmarbeitsverordnung / Verordnung über Sicherheit und Gesundheitsschutz bei der Arbeit an Bildschirmgeräten vom 4. Dezember 1996, zuletzt geändert durch Verordnung vom 31. Oktober 2006) (1996)
6. ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs). International Organization for Standardization / TC 159/SC 4 - Ergonomics of human-system interaction (1998)
7. ISO 9355: Ergonomic requirements for the design of displays and control actuators. International Organization for Standardization / TC 159/SC 4 - Ergonomics of human-system interaction (1999)
8. ISO 14915: Software ergonomics for multimedia user interfaces. International Organization for Standardization / TC 159/SC 4 - Ergonomics of human-system interaction (2002)
9. Abele, P., Jurtienne, J., Prümper, J.: Usability Management bei SAP-Projekten. In: Grundlagen, Vorgehen, Methoden, Vieweg, Wiesbaden (2007)
10. Gesellschaft Arbeit und Ergonomie - online e.V., `http://www.ergo-online.de/site.aspx?url=html/software/verfahren_zur_beurteilung_der/fragebogen_isonorm_online.htm`
11. Prümper, J.: Der Benutzungsfragebogen Isonorm 9241 / 10. Ergebnisse zu Reliabilität und Validität. In: Liskowsky, R. (ed.) Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung, Software Ergonomie, Teubner, Stuttgart, pp. 253–262 (1997)
12. ISO 9241/11: Ergonomic requirements for office work with visual display terminals (VDTs). Part 11: Guidance on usability. International Organization for Standardization / TC 159/SC 4 - Ergonomics of human-system interaction (1998)
13. Liebstückel, K.: SAP Enterprise Asset Management. SAP Press, Maryland (2009)

ISO Usability Standards and Enterprise Software: A Management Perspective

Clemens Lutsch

Microsoft Deutschland GmbH
Konrad-Zuse-Str. 1
D-85716 Unterschleissheim
clemensl@microsoft.com

Abstract. The introduction and manifestation of software quality according to ISO 9241 series in the enterprise ecosystem depends on many factors. Most of them are not technical ones but merely organizational and process-dependent. They become therefore hard to assess for an enterprise organization. It is essential to unveil these uncertainties, to address them and to solve them in a most constructive way that replaces the "angst" inside an organization seeking compliance with ISO 9241 series standards where applicable.

Keywords: international standards, user experience, human-centred design process, usability, Enterprise Software.

1 Introduction

We are entering a land of light and shadow! Emerging terms and paradigms like "User Experience" or "consumerization of IT" prompt managers and IT strategists to have a serious look at guidelines and standards that enable the development of interactive systems that will fit the picture painted by those somewhat vague but nevertheless emotionally loaded concepts. New products, faster comparisons (in the age of the Web, the competition is one click away) and also a changed perception by users on what quality of software really is lead to a noticeable pressure to act and to endorse human-centred design principles in enterprise software.

On the other hand, the problems are huge: legacy software all around, changing things that have worked for, oh, so long.

But there is a way to make the change happen. One key is the understanding that during the phase of adopting or change management toward ISO 9241 compliance, the major task for an organization is to avoid the fear of the new, to vitalize a culture of constant change, to be open to new competencies and the added value of usability and user experience professionals. This paper summarizes learning from more than 3 years of project work with different enterprise organizations and independent enterprise software development companies.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 154–161, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 The Situation

2.1 Enterprise Organizations

Different industries and markets result in very different patterns of enterprise organizations.

The use of software leads to a structured workflow in the organization that adheres to the structure of the software itself. Changes that are made in the software are reflected in the organization as well. IT has shifted from an assistive role in enterprise productivity (e.g. writing letters, calculating numbers) to an essential role (e.g. order management, communication center) which means the enterprise organization would stop being active and/or productive without it.

Adding weight to the effect of an enterprise organization on software is the fact that in some regions, the ergonomic quality of the workplace and all work related systems are mandated by law. It is therefore an obvious need of an organization to establish quality in interactive systems that meet the requirements stated in these regulations.

Management. Challenged by a subject that is quite often very far away from other daily work of a manager, two things are likely to happen when it comes to "managing" usability engineering:

1. Absolutely no management at all ("Do the stuff by yourself"), or
2. The manager becomes the most outspoken member of the usability team (regardless that the manager may have a limited understanding of usability but tends to use a personal opinion as decision threshold).

This might be regarded as a black and white differentiation, but rest assured, it fits the picture quite nicely. By establishing human-centred design in an organization, management is faced with many challenges that are outlined below:

- Managers are often not subject matter experts: they are not in the position to evaluate the employee's work
- Managers are seldom familiar with deliverables of the human-centred design of interactive systems: they cannot facilitate the collaboration of different team members or teams in development of interactive systems
- Managers are sometimes not used to evidence-driven design: they are not familiar with corporate KPIs of human-centred design as part of both system performance management and personal assessment.

Establishing a human-centred design team means assigning headcount; headcount means visibility in an organization; visibility means influence. Human-centred design teams therefore are quite often highly visible and political groups in organizations and become a political asset in organization politics [1], [2].

Experts. Today's usability engineering professionals are specifically trained in usability engineering methods and design activities. They still come from various disciplines like human factors, design, psychology, ergonomics, anthropology, computer science or economic pedagogics.

The community organizes itself in professional associations like the Usability Professionals Association [3], the Information Architects Institute [4], or special interest groups on computer-human interaction of the ACM [5].

The German Chapter of the UPA has formulated a definition of a usability professional, which describes a person whose work focuses in one or more of the following areas:

- Analysis - ascertaining the contexts of use, deriving the requirements for use
- Design - conceptualizing the interaction between human and system, structuring and illustrating required information
- Verification and Evaluation - conducting inspection based evaluations (without the user) and usability tests (with the user)
- Process orchestration and use of methods – determining, introducing and conducting human-centred design processes [6].

2.2 ISO Usability Standards

Standards that Focus on Activities and Methods. The most recent standard on activities related to human-centred design of interactive systems is part 210 "Human-centred design of interactive systems" of the ISO 9241 series "Ergonomics of human system interaction" [7]. It replaced the predecessor ISO 13407, which was called "human-centred design process" [8]. By dropping the notion of "process" in the new title the editors underline the change that human-centred design "enriches" process activities, regardless which process philosophy is applied (V-model [9], Rational Unified Process [10] or Agile Development [11] etc.)

An upcoming item in the 9241 family is part 230 on methods of human-centred design. Since this work is under discussion and far from being final, it should be stated that the audience and also the benefit of this part 230 still needs clarification.

A new group of standards are the upcoming common industry formats on the different "information items" that are results of the various human-centred design activities [12]. The term "information item" is used to stress the fact that information gathered or created during human-centred design is interrelated, iterative and dynamic.

Standards that Focus on Design. A series of ISO standards are in place that deal with guidance of user interface design. Some of them are undergoing investigation and/or are to be replaced with updated versions (e.g. on form-based dialogues [13]). These standards are to be interpreted by usability experts in order to use them efficiently in the design process. They must be viewed in relation to existing design guidelines (style guides), technologies and system constraints.

3 Preconditions

In order to successfully establish the ISO 9241 standard in an organization that creates enterprise software some preconditions are essential:

- The organization must feature a goal-driven personal assessment: each employee and manager has to define goals that they are to achieve during the fiscal year of the organization. The manager has to adopt human-centred design goals as personal goals.

- Manager goals are to be inherited by the team. Every specific goal of every team member contributes to the general human-centred design-related goal of the management.
- Empowerment of the human-centred design team to influence the software design.

4 The 8

4.1 Fundamentals

As an addition to existing development processes, the human-centred design approach is activity-driven and has to adhere where appropriate to development processes.

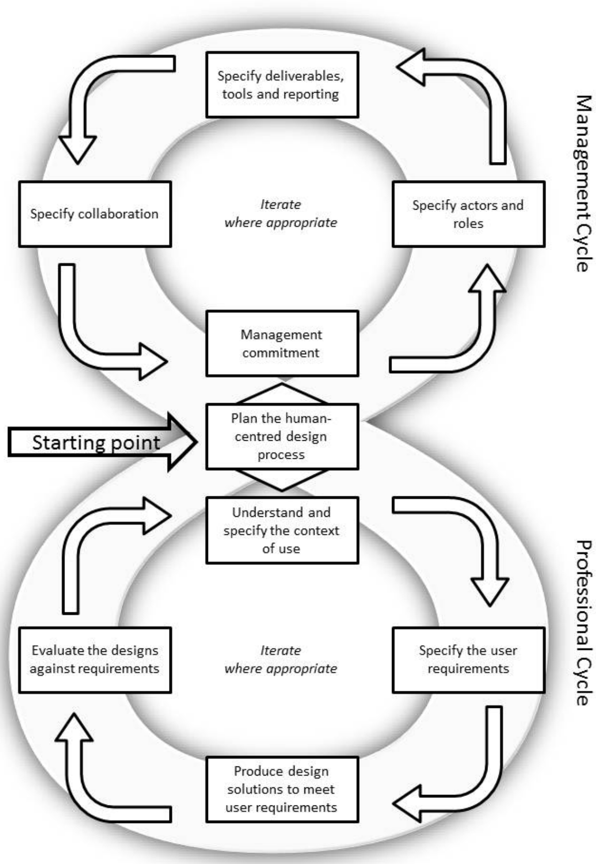


Fig. 1. The 8: Cycles of activities, a management cycle and a professional cycle; both contain essential activities for adopting ISO 9241 Series in an organization. Note that the professional cycle is adapted from ISO 9241-210 [7].

After discussions with managers, usability professionals and developers and after analysis of different organizations, a model emerged that can be used as a blueprint for activities that aim for adoption of ISO 9241 in the enterprise software context. Since this model (see Fig. 1) joins together two interlinked activity cycles, I chose the name “the 8” as title for this model.

Outlined in 9241-210 the design activities are visualized as an iterative cycle of activities, to be conducted by usability professionals (the professional cycle). To establish and maintain adoption of these ISO standards in an organization a second, iterative cycle has to be established that focuses on the stakeholders of the organization and their activities that are to happen synchronous to the human-centred design (the management cycle).

4.2 The Management Cycle

Overview. The starting point of the management cycle is inherited from ISO 9241-210: “Plan the human-centred design process”, which addresses the way to enrich the existing processes with human-centred design. In organizations this also adds weight to considerations of which stakeholders are affected by introducing human-centred design. These stakeholders could be other managers of software development departments, quality managers, project managers, product managers or the corporate design department.

The management activities in the “upper” part of the 8 are divided between: management commitment; specify actors and roles; specify deliverables, tools and reporting; and finally specify collaboration.

Management Commitment. Managers of teams that are affected by human-centred design (i.e., all teams engaged in planning, developing and testing enterprise software) are to formulate goals that address human-centred design as an essential part of work life. In addition, key performance indicators (KPIs) linked to aspects of human-centred design could also add weight to management and team commitment.

Specify Actors and Roles. Every person in an organization acts in a specified context and environment. In organizations with undefined work ecosystems, this person’s role is defined by the job title and the team this person is working in.

In organizations with a developed ecosystem (i.e., organizations with workflows, detailed job descriptions and evolved understanding of project management), the person is defined also by the competence and the different activities that are performed during a given workflow. Employers and also managers can be titled as actors, whereas this term refers to the person with a specific competence. Roles are linked to the activities that require detailed competencies. Depending on the capabilities of a person, one actor can embody more than one role.

Specify Deliverables, Tools and Reporting. In order to facilitate efficient adoption of human-centred design activities to ensure ISO compliant ergonomic qualities in enterprise software, tools and deliverables must adhere as well as possible to existing tools in software development. Nevertheless, some artifacts of human-centred design require other materials and workflows as well as tools. Depending on the different technologies used to develop the interactive system a huge variety of tools is available today. The multitude of tools and the mosaic of competencies in the usability /UX profession also generate a huge diversity in deliverables. It should be carefully considered to spend substantial time in specifying the tools and the deliverables. The organization must stress adherence to the agreed set of tools and deliverables to avoid disharmony in the team and finally the disconnect of human-centred design and system design. During the last few years we were able to collect supportive cases that using tools like Microsoft Expression Blend™ and Microsoft Visual Studio™[1] enables an organization to work more productively, because interface design and system design were working on the same project structure.

Specify Collaboration. After the team is set up, the next task in the organizational cycle is to establish collaboration between the different actors. By implementing ISO standards in enterprise software, the arrangement of a development team is changed. New activities must be performed, most likely by new actors. To avoid the team's reaction that "power" has been taken away (i.e., the power to decide upon the look-and-feel of the software), the organization must push the understanding of the benefit of human-centred design.

4.3 The Professional Cycle

Activities of this cycle are identical to those specified in ISO 9241-210. Those are: understand and specify the context of use; specify the user requirements; produce design solutions to meet user requirements; and evaluate design solutions against requirements.

5 Avoiding the Traps

5.1 The Statement

There are at least two big, dangerous and quite common traps in the set-up to introduce ISO 9241 in enterprise software.

5.2 Trap 1: The Costs Are "Already There"

This reflects the case when managers try to map human-centred design activities to existing personnel. Thinking that these employees "are already there" and that they "already did the user interface anyway", they should simply shift some work in order to embrace ISO 9241. The trap is that these employees may not be experienced or trained in usability engineering.

Example: in an enterprise organization we discussed this "already there" phenomenon in the light of the organization's reluctance to spend budget on an external user interface consultancy. The manager stated that four developers were also doing the user interface design. We were talking about four developers = 66k € per year spent on UI in the most inefficient way. For this kind of money the organization could acquire professional work from either an additional employee (a usability professional) or a specialist agency.

---

[1] Microsoft Expression Blend™ and Microsoft Visual Studio™ are registered Trademarks of the Microsoft Corporation.

5.3 Trap 2: We Do Everything But the Evaluation

The fear of the user's response and also the costs of evaluations are springing the trap that renders human-centred design almost useless when an software user interface evaluation is not conducted. Evaluation of design solutions is an essential part of the professional cycle. A decent choice of methods is required when setting up the human-centred design activities. A reflected mixture of different evaluation methods like usability testing, heuristic evaluation, questionnaires, log file analysis or focus groups could be a way to address different questions at different design stages.

6 Conclusion

After studying actual standards and discussing human-centred design activities within enterprise organizations, it should be noted that setting up ISO compliance may require substantial organizational changes. The multitude of aspects of software is mirrored in a kaleidoscope of management styles and attitudes. It is a strong belief of the author that without support in the organization, human-centred design will fail. Only superficial changes of the software in order to comply with ISO 9241 will not be sufficient: this will likely end in collaborative risks and problems and will affect the quality of the software. Nevertheless, a model for organizational activities that is synchronous to human-centred design activities for developing enterprise software has been developed, proposed and applied in some enterprise organizations. Further applications of "the 8" will validate its feasibility and will further the refinement of its contents and characteristics.

References

1. Lutsch, C., Petrovic, K.: Human-Centered Design als Unternehmensstrategie: Ein Arbeitsbericht. In: Brau, H., Diefenbach, S., Hassenzahl, M., Koller, F., Peissner, M., Röse, K. (eds.) Usability Professionals 2008, pp. 215–219. German Chapter der Usability Professionals Association e.V, Stuttgart (2008)
2. Petrovic, K., Lutsch, C., Weißenberger, U.: Das unentdeckte Land: User Experience in der Unternehmensstrategie. In: Brau, H., Diefenbach, S., Göring, K., Peissner, M., Petrovic, K. (eds.) Usability Professionals. German Chapter der Usability Professionals Accociation e.V. Stuttgart, pp. 82–85 (2010)
3. The Usability Professionals Association, UPA, `http://www.upassoc.org`
4. The Information Architects Institute, IAI, `http://www.iainstitute.org`
5. The Special Interest Group on Computer-Human Interaction of the ACM, SIGCHI, `http://www.sigchi.org`
6. Berufsfeld Usability/UX, German UPA (2010), `http://germanupa.de/german-upa/berufsfeld-usability-ux`
7. International Organization for Standardization: ISO 9241-210: Ergonomics of human-system interaction –Part 210: Human-centred design for interactive systems. ISO, Geneva (2010)
8. International Organization for Standardization: ISO 13407 (Withdrawn): Human-centred design processes for interactive systems. ISO, Geneva (1999)

9. The V-model, http://www.v-modell-xt.de
10. The Rational Unified Process RUP, http://www.ibm.com/rational
11. Agile development, http://www.agilealliance.org
12. International Organization for Standardization: ISO/IEC TR 25060: Systems and software engineering – Systems and software product Quality Requirements and Evaluation (SQuaRE) – Common Industry Format (CIF) for usability: General framework for usability-related information. ISO, Geneva (2010)
13. International Organization for Standardization: ISO/DIS 9241-143: Ergonomics of human-system interaction – Part 143: Form-based dialogues. ISO, Geneva (2010)

Guidelines on Website Design and Colour Selection for International Acceptance

Martin C. Maguire

Loughborough Design School, Loughborough University,
Ashby Road, Loughborough, Leics, LE11 3TU, UK
m.c.maguire@lboro.ac.uk

Abstract. Since websites are accessible from all parts of the globe and across many nationalities and cultures, the challenge for designers is how to make them suitable for international use. While it may not be possible to appeal to all users that may access a website, it is possible to maximize its acceptability and effectiveness internationally. This paper describes published guidelines for achieving a certain degree of international and cross-cultural acceptance without necessarily requiring a large amount of work and effort. It also describes a study conducted on colour association with a sample of users from different countries, and how the results might assist designers.

Keywords: website design, international use, cross-cultural use, colour, usability.

1 Introduction

In an increasingly interconnected world, web designers who want their sites to be used within different countries and across cultures need to consider a range of design issues. This is particularly true for commercial organisations hoping to sell products worldwide. This paper reports on published advice for web design to help make websites effective and acceptable internationally. It addresses websites in general as well as aspects specific to e-commerce sites. It also describes a study conducted by the author on the use of colour in websites based on a sample of users from different countries.

2 General Strategy

When devising a strategy for internationalization, it is important to decide whether nor not to have multiple foreign web sites and how to manage them. One approach is to set up a single site and to adapt it centrally for different local countries. This helps maintain a consistent brand image across different markets. Companies such as the Dell Inc. have managed in each of its local sites to convey a common brand in selling computers. However mistakes can be made such as linguistic errors in the text used on the site and, for a commercial site, possibly selling products that are inappropriate in certain countries or cultures.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 162–171, 2011.
© Springer-Verlag Berlin Heidelberg 2011

The alternative method is to hire a team to handle the site in the local country. The normally requires the oragnisation to employ someone in the host country to take responsibility for and to handle enquiries received via it. Such a site is much more likely to meet local requirements but can lead to problems for companies trying to convey a consistent brand message.

3 Language

An important issue is whether an international company or organisation should have its entire websites translated into different languages. There is a strong argument for this as approximately two billion people across the world currently access the web (a figure that will grow rapidly). Although English may be seen as a common language for website content, the language used by most people in the world is Chinese Mandarin, followed by English and Spanish. Only one quarter of the world's population speaks English, the vast majority of which speak it as a second language [1].

While large organisations may consider creating multiple versions of their websites in different languages, this may not be possible for smaller organisations with limited budgets. For a European audience, a compromise might be to offer versions of the site in English, French, German, Spanish and Italian. Another option is for those with limited resources is to consider translating an introductory or 'welcome' page into several of the most important languages that the site is intended for.

Online translation is improving rapidly so that users now have the option to paste sections of the website into a translation package to be able to read the text in their own language. To assist with this, the website may point to an online webpage translation service such as Yahoo Babel Fish. A more convenient method for the user is to add a translation widget to the website such as from Google Translate or Microsoft Live [2]. This allows webpages to be translated dynamically with the user simply selecting the language they require. The conversions are based on machine translation but services are now being provided with the option for the designer to select human translations for sections of text where this has been carried out and the translation stored.

Websites that aim to be accepted internationally should avoid phrases that are colloquialisms or slang only known to the local country. The use of the term 'hardback' and 'paperback' commonly used to distinguish books in English may not have a clear meaning in other European countries. For an e-commerce site the terms 'store', 'aisle' and 'checkout' may be familiar in countries with established supermarkets but may be unclear elsewhere.

Spellings of words such as 'color'/'colour' or 'license'/'licence' differ across countries reflecting the minor but possibly important differences between American and British English. Getting the spelling right for the correct market may be important for some customers and users; see [3].

4 Ensuring Correct Characters

One problem for web designers intending to reach international audiences is the need to handle international character sets. It is thus recommended to design the website using Unicode. UTF-8 is a variable-length character encoding for Unicode that is compatible with over 90 scripts (written languages) and 100,000 characters. Unicode has been adopted by most of the large IT organisations, including Apple, Microsoft, IBM, Oracle, and is compatible with all the common browsers and operating systems. Using UTF-8 ensures that characters can be used from many non-English alphabets and be displayed properly at the user's end [1].

5 Forms

In developing an online form, the likely target audience or market must be kept in mind. The information collected from overseas respondents or customers should be kept separate from domestic information so that they both can be seen in context and handled efficiently.

If the website includes an order form or requires input of address details, then it should be compatible with the formats of international addresses, titles, surnames, telephone numbers, postal/zip codes, etc. Care should be taken as to which fields are made obligatory. 'Province', 'county' or 'state' should not be obligatory fields, for example, as many countries do not use these terms.

Titles for people are important in many countries so forms should include such a field or allow a title to be entered with the name. Forms should allow enough space for complicated street addresses, for a large number of digits for telephone and fax numbers. There should of course be a country field. The postal or zip code field should also be able to handle letters as well as numbers. The surname field should be capable of handling multiple last names [4] (e.g. 'Martha Lane Fox') or hyphenated names ('Chou Wen-chung').

6 Layout and Context

Another issue of importance for international website design are graphic features such as pictures, drawings, and illustrations. If graphics only make sense when read from left to right, then there might be a problem in countries where the population read from right to left. With some graphics, e.g. representing a sequence of actions, this could change their meaning completely. Similarly, organisations should be sensitive to the graphics they use and how they will be received in different countries. International firms should use icons, symbols and design features on their web sites that will not be confusing to others. The popular sign for 'OK' in some countries and in the diving world (the thumb and forefinger brought together in a circle with the other three fingers up in the air) has a derogatory meaning in other parts of the world [4].

Some authors suggest that users in some countries are happy to view a higher density of characters and graphics on webpages. This may be true but web styles in different countries will evolve over time so it is perhaps risky to make generalizations about the amount of characters that users find acceptable. In general it is preferable to concentrate on a single usable layout rather than artificially enrich a page for a particular country or culture.

Context and culture is also an important topic in website design and its effect on user behavior when using websites [5]. It is widely believed that people from 'high context' cultures are particularly interested in background information on a website such as about the company, the people running it, its business motto, history, place in society, and its relationship to themselves. They are also thought to be more accepting of diverse layouts, imagery and multimedia. In 'low context' cultures it is suggested that the emphasis needs to be on the most important information needed to use the site and how it can be clearly expressed. User preferences may also be for simpler and more minimal layouts [6]. In the author's view, consideration of both levels of cultural context should lead to the development a website that satisfies both viewpoints.

7 Credit Cards and Alternative Payment Mechanisms

While credit cards may be popular in many countries for online payment purposes, they are less popular elsewhere. Some countries may use debit cards more often, cheques, or online payment services such as World Pay or PayPal. In Japan, for example, local 7-Eleven stores are used by consumers to pay for many services including utilities and many other types of bills. The organisation should try and allow for the most popular methods of payment for the audience it is targeting. Possible methods of payment should also be described upfront rather than only being shown after the user has gone through a long process of finding and selecting the items they wish to buy. Also customers in a particular country will be sensitive to having to pay commission on the method of payment that is most popular in that country if other methods are commission free.

If the organisation were targeting consumers in Japan, for example, there is less interest in credit card payment than in other places. Instead the organisation could enter into an agreement with 7 -Eleven for them to accept payment on the firm's behalf and their commission would need to be calculated into the firm's selling price [4]. In addition, the international firm should always offer an alternative way of paying for a purchase such as provide its bank details or suggesting that the customer contact them to discuss alternative payment methods.

8 Responses to Enquiries

Handling emails from visitors is another issue for website management. The organisation might wish to let potential customers know that they can communicate with them in their language of choice. But they will then need to have a translation service on hand locally to translate these e-mails into the local language of the organisation which can be expensive if they are receiving a large number of foreign-language e-mails daily. International organisations need to ensure that any emailed enquiries they are answered and dealt with as quickly as possible. Potential customers may send enquiries from different time zones and yet will still expect a rapid reply. The organisation may need to set up an auto response service that sends an immediate reply to the enquirer thanking them for the enquiry and informing them that they will receive a full reply within 24 or 48 hours [4]. Having local customer service staff is equally important as having a localized website so that when a user or customer picks up the phone, they are dealing with someone local.

9 Time and Date and Units

Some countries use a 24-hour clock (written, for example, as 13:20 or 13h20), while others use the 12-hour clock (written as 1.20pm). The format should make it clear which one it is using. Indication of local time in the home website company is also important if this affects the time when a response can be provided to a user query or order.

Countries such as the UK write the date as either day/month/year (e.g. 24/06/2011) while others e.g. America format it as year/month/day (written as 2011/06/24). This can be confusing when the date is between 1 and 12 which may be misinterpreted as the month. One way to overcome this is to write the date with the month represented in letters e.g. '24 June 2011' which helps to prevent misunderstandings. However the ISO standard 8601 [7] specifies that the international standard date notation is: YYYY-MM-DD and for time is HH:MM:SS.

For e-commerce websites, different currencies can cause problems and users are often uncomfortable if the cost of an item is shown in, say, pounds sterling, dollars, euros or yen, if this is not their local currency. Pricing of goods and service in a local currency is the most preferred or alternatively providing a currency calculator on its web site. Web sites selling goods also need to make sure that they are kept up to date with the latest prices, tax rates, product specifications or any other information changes, to avoid user frustration. The cost of delivery to the user's local address should also be specified.

In many countries the metric system (metres, grams and litres) is used, but in the US the imperial system (feet and inches, pounds and ounces, gallons and pints) is employed. In the UK, metric is used formally but imperial is often used informally. International firms might consider offering both measures for ease of reference.

10 Use of Colour

Research has shown that particular colours can mean different things in different countries [8] and cultures where a cultural group may be ethnic, religious, social, professional, sports, activity based, etc. The colours used to represent a country, such as those appearing on the flag, can have very significant meaning within that country e.g. the red, white and blue of the Stars and Stripes in the USA, the Union Jack in the UK or the Tricolour of France. Other colours are strongly associated with a country although not components of their flag such as the green and gold of Australia. At another level (both local and internationally) the colours of a football team have strong meanings to their supporters, as reflected in football club websites.

Colours also have a symbolism for people through cultural references. Depending upon the culture, colours can have quite different meanings. They can stimulate an emotional reaction so that in Eastern countries for instance, red is a very positive colour and is worn by brides and evokes happiness and prosperity. In Western countries it has a different connotation and is an indicator of excitement, danger, and passion.

The web designer may wish to stimulate positive reactions or at least avoid negative reaction. For example, while in the West white is the colour of weddings, in the East, it is the colour of mourning so would be inappropriate for a wedding site. [9]. Tables of colour associations (e.g. [10]) for different countries or regions can be useful can be a useful guide for colour selection and importantly for avoiding negative associations.

For some colours, the origins of the symbolism are clear –green's association with nature, plants, new growth and the environment while the unspoiled colour white represents purity. Also, in general, cool colours such as blue and green tend to be associated with calmness and relaxation, while warm colours such as red and orange often have a more emotional meaning.

Of course the topic or application of the website also influences the choice of colour and can receive fairly wide cultural or international acceptance. So, for instance, blue is often used in banking websites since in Eastern countries as it is seen as an indicator of wealth, while in the West it is a conservative colour traditionally denoting stability and the careful handling of money.

A study was performed of preferences for assigning colours to tabs for logos or 'favicons' in a web browser. This found that colour assignment to each tab based on (i) the primary colour of the webpage, (ii) the website logo or favicon, (iii) personal preference and (iv) emotional experience and website expression, were all helpful for creating an efficient performance and pleasant experience for both European and Asian users [11].

Overtime each national or cultural group have developed their own meanings of colours although there can be commonality between traditional cultural meanings for certain colours. The question arises then of how to select a major colour when designing a website for a particular cultural or national group? The following steps might be considered [12]:

1. If there is a specific colour that a region or culture associates with the topic of the website (e.g. blue representing heaven and spirituality, pink indicating trust) then it could be a good choice of colour for that topic.
2. If there is no specific colour that is suitable for the topic, consider the feeling that website is intending to convey (e.g. excitement, creativity, reliability, wealth, trust) and try to choose a colour that stimulates it for that particular region or culture.
3. Avoid choosing a colour that contradicts the topic area of the website or may be likely to evoke a negative reaction in the audience.

It is also helpful to consider the level of sensitivity of the website topic to a specific choice of colour. Table 1 below suggests possible sensitivity levels for certain topics.

Table 1. Potential user sensitivity to website colour for certain topics

Low sensitivity	Medium sensitivity	High sensitivity
Social sites	Business/financial sites	Religious sites
Art/creativity sites	Shopping/holiday sites	Traditions & customs
Music sites	Health/scientific sites	Political sites
News or media sites	Academic sites	Sports fan sites

Thus for a website related to art and creativity, a radical or unexpected choice of main colour may be helpful to make a strong impact and users are likely to be open this. For business or holiday sites, users are likely to be more sensitive to appropriate choice of colour, while for certain sites (e.g. religious or cultural tradition) the user is likely to be highly sensitive to colour choice.

A small survey was conducted by the author with a international group of 55 web users from 14 countries to assess the colours they associated with different types of website. The sample included people from Brazil, Bulgaria, China, France, Germany, Jordan, Netherlands, Sri Lanka, Spain, Switzerland, Ukraine, Venezuela, UK and USA Each person was asked to consider 8 different colours in turn and to name the types of website (one or more) they associated with each colour. Figure 1 shows the results of this study.

In the first part of the survey each person was asked to consider nine different categories of website and to select from a list which colour or colours they associated it with. The results in Figure 1 show that colour associations for financial and art/music websites tend towards the blue and purple end of the spectrum while for shopping and holiday/vacation websites, they tend more towards the red, orange and yellow. The subjects selected a higher number of colours for children's websites indicating that they either felt a wider range of colours were appropriate or that they expected them to be multi-coloured.

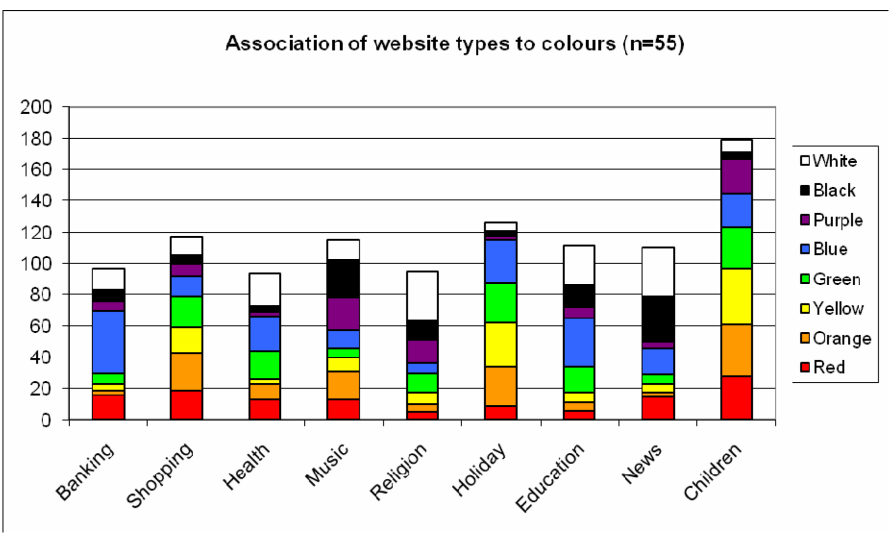


Fig. 1. Website types and colour association. Each user presented with 9 website types and asked to state which colours (multiple choice) they associated with each type. Colour and section length of bar indicates the number of participants naming a particular association.

In a second part of the survey, participants were asked to consider a selection of 8 colours and to say which website types they associated with each. The results are shown in Figure 2. Bright colours such as yellow and orange were again associated quite strongly with holiday/vacation and shopping sites but also with travel-transport sites. Blue and red were quite strongly associated with money or banking sites. Black was associated with art and music sites while black and white was often associated with news and information websites. As well as being representative of holiday and shopping sites, green seemed to be an indicator of nature or the environment, food-forming, education and health.

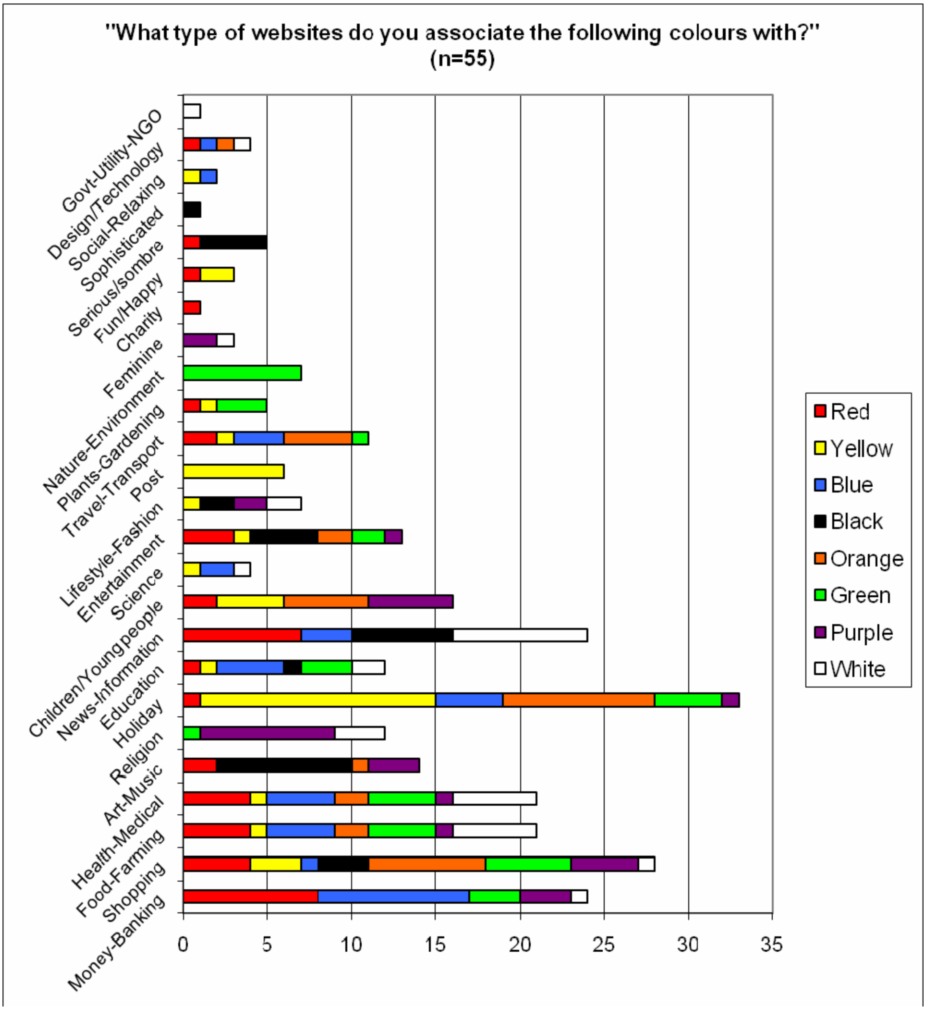


Fig. 2. Colours associations and website types. User presented with colours and asked to state website type (free choice) they associated each with it. Colour and section length of bar indicates the number of participants naming a particular association.

It is likely that some respondents were influenced by branding colours used by certain organisations. For example yellow was associated by 6 people with postal services as is used in many European countries e.g. by Deutsche Post in Germany.

The results of the survey may give an indication to designers to help choose colours to meet user expectations or at least avoid creating a 'culture shock'. Note that tailoring web colours to particular regions need not create stark differences. Use of subtly different colours and designs for websites in different countries can also be effective if used appropriately. Of course, a larger survey would be needed to validate these findings and perhaps show differences between cultures and countries.

PS: In the survey each person was also asked to name their favourite colour. Blue was the most popular colour named by 24 people out of the 55 (14 males and 10 females), followed by orange (12 people), red (10 people), green (9 people), black and purple (6 people each) and yellow (5 people).

11 Conclusions

While these guidelines may offer some guidance to designers of websites for an international audience, it should be remembered that tastes and preferences in different countries will continue to develop and change over time. What is desirable for these audiences should therefore be assessed on a continuing basis. Globalisation has already led to major company brands and their associated websites being accepted across the world. However there will always be the need for websites to appeal to local preferences and to be sensitive to local cultures, resulting in a positive impact on local audiences. Even though a website cannot attempt to be fully integrated with the conventions and needs in every country or culture, this approach can go a long way towards websites being accepted internationally.

References

1. Woods, D.: Perspectives on Cross-Cultural Web Design, March 12 (2010), http://www.dave-woods.co.uk/index.php/perspectives-on-cross-cultural-web-design/
2. Osmani, A.: Web design for every culture, June 20 (2010), http://addyosmani.com/blog/web-design-for-every-culture/
3. Bond, K.: Differences between British, Canadian and American Spelling, http://www3.telus.net/linguisticsissues/BritishCanadianAmerican.htm
4. Export help, http://www.exporthelp.co.za/marketing/internationalise_site.html
5. Würtz, E.: A cross-cultural analysis of websites from high-context cultures and low-context cultures. Journal of Computer-Mediated Communication 11(1), Article 13 (2005)
6. Arno, C.: Web Design across Culture and Context, http://www.webdesignideas.org/2010/03/09/web-design-across-culture-and-context

7. ISO 8601:2004: Data elements and interchange formats - Information interchange - Representation of dates and times, Geneva, Switzerland (2004)
8. Kitao, K. and Kitao, S. K.: A study of color association differences between Americans and Japanese. Human Communications Studies 13, 59–75, (1986), http://eric.ed.gov/PDFS/ED273134.pdf
9. Kyrnin, J.: Colour symbolism - Using colors to convey meaning may backfire on a web site, http://webdesign.about.com/od/color/a/aa072604.htm
10. Kyrnin, J.: Color symbolism chart by culture, http://webdesign.about.com/od/color/a/bl_colorculture.htm
11. Fan, Y.: Comparison of Colour background Preferences Applied to Browser Tab design between UK Students and Chinese Students, MSc project, Department of Human Sciences (now part of the Loughborough Design School), Loughborough University (2009)
12. Maguire, M.: The use of colour in intercultural website design, Gesellschaft für Arbeitswissenschaft (GfA - Society for Industrial Science), neue Arbeits- und Lebenswelten gestalten (creating new work and living environments), Darmstadt, Germany, March 24 – 26, pp. 877–880 (2010)

User-Experience for Personal Sustainability Software: Determining Design Philosophy and Principles

Aaron Marcus, Jennifer Dumpert, and Laurie Wigham

Aaron Marcus and Associates, Inc., 1196 Euclid Avenue, Suite 1F,
Berkeley, CA, 94708 USA
{Aaron.Marcus,JenniferDumpert,Laurie.Wigham}@AMandA.com
www.AMandA.com

Abstract. Business developers worldwide seek to develop sustainability software with a user experience that provides usability, usefulness, and appeal. This paper describes the research and analysis that led to a design philosophy and a set of principles that were then "tested out" in a series of short prototype applications intended to make more compelling and engaging business applications incorporating sustainability and taking advantage of people's interests, expertise, and experience with sustainability.

Keywords: Business, design, development, enterprise software, management, user interface, sustainability, user experience.

1 Introduction

User-experience development (UXD) and sustainability are both of increasing concern to developers of business-oriented applications worldwide. This project describes the activities undertaken to determine a UX philosophy and set of design principles that could be explored through conceptual prototypes, then evaluated, before attempting to apply them to actual product development..

In August, 2010, SAP engaged the authors' firm to envision possible products in the space of personal sustainability. AM+A was selected in part because of its prior demonstration of a consumer-oriented mobile conceptual prototype called the Green Machine, which combined information design/visualization and persuasion design to change people's behavior regarding sustainability [Marcus, 2009, Jean and Marcus, 2010].

AM+A created four initial concepts and then merged them into a final concept that was presented to SAP on 4 October 2010.

2 User-Centered Design Process

In order to explore what motivates people to make changes around sustainability, and to evaluate current trends in the market space, AM+A conducted research in knowledge visualization and conducted competitive analysis, content analysis, and user research. AM+A then created design principles based on that research. With

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 172–177, 2011.
© Springer-Verlag Berlin Heidelberg 2011

those principles as basis, AM+A conceived four possible personal sustainability products, Green Buddy, Gas Gauge, Credible Edibles, and Green Jeeves. Working in partnership with SAP, AM+A then combined several features from Green Buddy and Green Jeeves to create a final sustainability product.

2.1 Market Research

AM+A surveyed current sustainability-related applications for mobile devices, focusing on the iPhone, and the Web. AM+A analyzed the results to observe trends, effective solutions, and current industry best practices, then to discover which images and design styles had become “stale” from over-use/over-exposure and were to be avoided.

Applications that provided clear, easy-to-follow advice, such as the GoodGuide and Seafood Watch, seemed especially popular.

Websites. Websites can provide more detailed and complex tools than mobile phones applications. Users seemed more likely to invest more time and thought, and might be willing to log in and save results over multiple sessions. Some of the Websites surveyed were so complex and hard to understand that users seemed likely to be unwilling to struggle through the process. Others were so over-simplified that their results seemed meaningless. Very few sites appeared to find the sweet spot that balanced relevance with ease of use. AM+A also looked at Websites that had the following specialized tools:

Sutainability ratings for consumers
Calculator tools for measuring personal energy usage
Other calculators, such as water usage
Dashboards and meters

AM+A found that the market was “flooded” with personal carbon-footprint calculators. Many of these calculators seemed difficult to use. Many Websites featured these tools as a way to convince users to purchase carbon offsets or carbon management solutions from the site,. Somewhat surprisingly, most of the carbon-footprint Websites featured a palette of blues, unlike the greens which predominated in the iPhone applications. This color palette may be selected to connote an image of blue skies uncontaminated by greenhouse gasses, or cool colors unaffected by global warming.

Knowledge Visualization. AM+A collected examples of information graphics and data visualization, then presented them to the client to stimulate thinking about similar techniques that could be carried out with SAP’s date.

2.2 User Research

Over the course of the project, AM_A conducted user research with a total of 108 participants in order to gather information about concerns around personal sustainability and to glean information about what motivates people to make behavior changes.

Content Analysis. AM+A's Content Analysis had two major objectives:

- Determine what particular sustainability issues stand at the front of popular consciousness
- Examine language used to describe sustainability issues

Key Sustainabilty Issues. To explore prominent issues, AM+A generated tag clouds from leading news sites to determine frequently used and mostly highly weighted words. AM+A first generated a tag cloud of 200 words for each site, then reviewed the cloud for each site and selected words relevant to sustainability concerns. AM+A then returned to the news site and searched articles using each word, then noted what the topic was of relevant articles. Using this method, AM+A filtered the most frequently used words around sustainability, then identified what subset of sustainability concerns appeared most frequently in popular news media.

Sustainabiity Language Use. To learn more about language, AM+A generated tag clouds from leading environment organizations, based on membership. By examining and cross-referencing frequently used words and phrases, AM+A generated a short glossary of sustainability nomenclature.

"Man on the Street" Interviews. This method uses a "buttonholing" approach to gathering subjects. The researcher stands in a well-populated area and recruits potential interview subjects. The objective of this approach is to "take the pulse" of everyday people in order to discover their main concerns, gain insight into what language people naturally use, and to gauge their aversions. For this user research method, AM+A sought specifically to gather un-premeditated responses. AM+A recorded audio of interviews to include as part of its final presentation to the client. Over the course of three days, AM+A interviewed 18 subjects, both at work and shopping in two different supermarkets (one in Berkeley, California, and one in Walnut Creek, California) and in a mall (in Walnut Creek). These sites represent different demographics: liberal and more conservative.

Focus Groups. AM+A also conducted two focus groups on-site at SAP with SAP employees as participants. The objective of the focus group was to promote open discussion among participants who had had time to consider the topic at length. Participants were encouraged to offer their opinions and to develop their ideas through conversation. The interaction and development of opinions throughout the session provided information about what changed and affected the perceptions of participants around sustainability, and about what ideas maintained strength and constancy when challenged and explored.

AM+A conducted two 90-minute focus groups over the course of two days in two sessions, with a total of 13 participants. These participants were asked to respond to the same five questions that were used in "man-on-the-street" interviews; however, for focus groups, participants were given the questions beforehand. At the end of the sessions, participants split into pairs and were given a video camera. They conducted brief interviews *with each other* about key points in the session. Short video segments of participants' responses were used in the final presentation to the client.

Global Survey. To expand user-research feedback across different cultures, AM+A created a Web-based survey provided to SAP employees worldwide using the same five questions that formed the foundation of its prior user research. A total of 76 participants, from 17 countries, responded. These responses were then compared with other user research to create trends in user response.

Synthesis Session. In order to surface user-research trends and organize these trends into categories and priorities, AM+A conducted a synthesis session with key SAP stakeholders. Using post-it notes (a standards UCD technique), the team named, organized, categorized, and prioritized user research findings.

2.3 Results: Key Trends

Based on the previously described user research methods, the most high-level trends observed in sustainability were the following:

- Lifestyle is a primary concern.
  - Any solution must fit into potential users' current lifestyles)
- People have a lot of confusion about what the right thing to do is, and they share a universal sense that it is important to do *something*, although they are unsure of what that something should be
- The topics most often in the media and cited by people as most important are these:
  - Energy (oil, fuel, cars)
  - Environment
  - Wasteful consumer culture
  - Food
- The actions people actually take most often around sustainability are these:
  - Recycle
  - Take public transit, bike, walk
  - Eat local organic food, less meat
  - Buy energy-saving things: cars, meters, solar panels, *etc.*
- Users distrust sustainability efforts in business, considering it a marketing ploy, or "green washing" to capitalize on a trend or fad to make profit
- People want guidance from trusted sources. Distrust of green marketing leads to distrust of companies. The meaning of labels and terms is often unclear, e.g., "organic", "free range", "natural"
- The things that motivate people to actually make changes in their lives around sustainability involve both "carrots and sticks."
- Carrots include:
  - Infrastructure, e.g., recycling bins
  - Competition and "cool" points
  - Financial benefit
  - Benefit for children, the next generation
- The primary stick is this:
  - Forced change is acceptable within boundaries, *e.g.*, making people pay for plastic bags, or a company no longer providing plastic water bottles

- The most effective influences that lead people to take action are:
  - Community, trusted friend, teacher
  - Company: employees of SAP cited SAP mostly positively
  - News, online sources, popular media, *e.g.*, the “Al Gore movie”, *Omnivore’s Dilemma*, or striking images
  - Negative perception of “preachy” approaches, “eco-nags”, or “extremists”
- User research also uncovered key findings around language:
  - Most users define “sustainability” in ecological terms, *i.e.*, maintaining resources
  - The term “green” is sometimes viewed as a marketing tagline
  - The term “Eco-” is overused in common parlance, even though it occurs surprisingly infrequently in news sources
  - Overall, “sustainability” and “sustainable” still maintain the most positive associations

Based on these trends, AM+A developed user-centered design principles to guide their initial concept designs.

3 Design Concepts

Following the lessons learned from its research, the key trends, and its design-principles formulation, AM+A proposed four initial concepts for product directions that would combine sustainability with business applications.

For the final concept, AM+A combined Green Jeeves and Green Buddy into a single application.

AM+A created several sample screens and wire frames to illustrate this application and presented the final concept to the client.

4 Conclusions

In general, the client response was favorable in regard to the usability and usefulness of the trends identified, the design principles formulated, and the effectiveness of their embodiment in the final concept design.

Acknowledgments. The authors acknowledge the assistance Ms. Janaki Kumar, Mr. Garrett Miller, and Dr. Dan Rosenberg of SAP for their guidance and support in the preparation of this paper. The authors thank, also, numerous SAP employees for their participation in discussions and survey responses

References

1. Jean, J., Marcus, A.: The Green Machine: Going Green at Home. User Experience (UX) 8(4), 20–22 (2009)
2. Marcus, A.: User-Centered Design in the Enterprise. Fast Forward Column. Interactions 13(1), 18–23 (2005)

3. Marcus, A.: User Interface Design's Return on Investment: Examples and Statistics. In: Mayhew, D.J. (ed.) Bias, Cost-Justifying Usability, 2nd edn., ch. 2. Elsevier, San Francisco (2005)
4. Marcus, A., Jean, J.: Green Machine: Designing Mobile Information Displays to Encourage Energy Conservation. Information Design Journal 17(3), 233–243 (2010)

ISO Standards and Enterprise Software: A Case Study Using SUMI and SUS in an International Sale

Joyce Ohgi and Jeff Sauro

Oracle USA, 300 Oracle Parkway MS3op343, Redwood Shores, California 94065 USA
joyce.etsuko.ohgi@oracle.com

Abstract. Usability standards play an important role in Europe, where companies are legally obligated to ensure that the workplace meets software ergonomic standards. Therefore, usability can be a critical factor in a company's decision about which enterprise software solution to implement. In this paper, we provide a real-world case study where meeting ISO standards and providing usability data to a customer heavily influenced a competitive sales deal.

Keywords: ISO, SUMI, SUS, international usability testing, enterprise software, user experience, software procurement.

1 Introduction

A prospective European customer included usability metrics as a requirement in its request for tender (RFT) document. The customer wanted to evaluate Oracle E-Business Suite 12 (EBS 12) applications along with other enterprise solutions. The customer wanted a clear and objective measurement of user perceptions regarding the usability of the software and how usability affected a user's ability to accomplish daily work. The customer requested Software Usability Measurement Inventory (SUMI) scores from all 10 of the enterprise applications that it was considering. The customer selected SUMI [1] because SUMI is referenced in ISO 9126 [2] and ISO 9241 [3].

The Oracle sales consultant working on the RFT reached out to the Oracle Applications User Experience team to provide the requested usability data. At the time, the Oracle user experience team had System Usability Scale (SUS) data for seven of the 10 applications under consideration and a SUMI score for one application, and could satisfactorily provide the correlation between SUS and SUMI data to meet the RFT requirement. Oracle lacked data for two of the applications, however, and needed to collect SUMI data for these two applications in less than two weeks to meet the RFT deadline.

The Oracle sales consultant identified an existing customer (located in the same country as the prospective customer) who was willing to provide qualified end users to participate in a study to collect the necessary data. Oracle flew a usability engineer to this customer site to conduct the study and collect SUMI data on the two remaining Oracle EBS R12 applications and to create a report in time to meet the RFT deadline.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 178–184, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Procedure

2.1 Materials

To collect the usability data, we needed to rapidly create, validate, and translate all of the following key materials:

- **User profile.** We needed to identify and validate the appropriate user profile for the applications being tested. The user profile ensured that we were conducting the study with the appropriate, qualified users. Oracle's existing repository of user profiles enabled us to quickly select the relevant user profile and obtain the presales team's validation.
- **Hosted test environment.** We needed a stable environment in which to run the test sessions, ensuring a consistent experience across all 10 users. Oracle's internal technical support team set up the environment for the test.
- **Task list of key product flows.** We needed a task list of the key product flows in the two enterprise applications being tested. We used two primary criteria to determine the key tasks: (1) The tasks needed to reflect the most common actions that a typical end user would accomplish within the applications, and (2) the average test session length needed to be around 1.5 hours, which is fairly representative for the usability tests conducted at Oracle. The task list that we created included six key tasks, with scenarios and sample data that participants needed to complete each of these tasks. Participants used this task list during the test sessions.
- **Test plan.** We needed a test plan for the usability engineer. The test plan that we created included: the task list; the expected starting point, click path, and ending point for each task; and the sample data that needed to be seeded in the test environment for each task.
- **Introductory screener and script.** To ensure that each participant met the criteria listed in the user profile, we created a document with a short list of questions asking about the participants' backgrounds and job roles. We also included a written introduction to the study in the users' native language. Although all of the users spoke both their language and English, we wanted to be certain that the users received the instructions in their native language, as the usability engineer spoke only English.
- **SUMI and SUS scores.** Although the RFT required only SUMI data for Oracle EBS Release 12 applications, we decided to collect both SUMI and SUS scores. Depending on the results, we knew that having both scores could give our statistical analyst further data to illustrate the correlation between SUMI and SUS. In addition, we opted to collect SUMI and SUS scores for both Oracle EBS 11 release 11.5.10 and Oracle EBS 12 applications. Again, depending on the results, we knew that these scores could show increased user satisfaction with the newer Oracle EBS 12 applications over the Oracle EBS 11 release 11.5.10 applications, demonstrating an improvement in usability over time with Oracle applications.

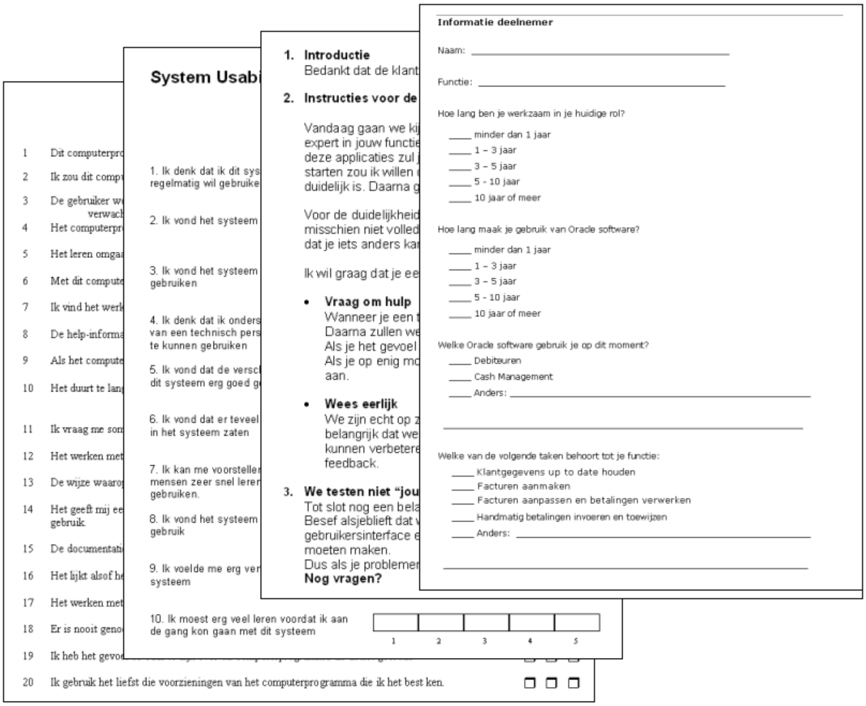


Fig. 1. Sample of test materials translated into a user's native language

2.2 Method

From an existing Oracle customer located in the same region as the prospective customer, we identified 10 users who fit the identified user profile. These users were also familiar with Oracle EBS 11 release 11.5.10 applications.

The usability engineer flew out to the customer site to conduct the sessions in person. Each test session involved one participant and was limited to 1.5 hours. Each participant was greeted by the usability engineer and was oriented to the test procedure in his or her native language. The usability engineer explained how the evaluation would proceed. The participant read each task, attempted each task, and indicated when he or she was finished with each task. The user was given six tasks to complete with the Oracle EBS 12 applications. The user worked to the best of his or her ability to complete these tasks. At the end of a session, each user completed SUS and SUMI surveys (also translated into his or her native language) on the Oracle EBS 12 applications.

3 Satisfaction Measures Overview

3.1 SUMI

SUMI is an inventory questionnaire specifically designed to measure user satisfaction with products. SUMI data has been collected from hundreds of thousands of participants. This data enables researchers to compare industry norms for satisfaction with usability on five main factors: efficiency, affect, helpfulness, control, and learning. In addition, there is a global factor that can be used when a single, global user-derived value is needed. Oracle typically uses this SUMI questionnaire as part of its usability methodology on software going through the usability labs in a post-release form.

3.2 SUS

The SUS questionnaire is another satisfaction measurement tool for computer systems [4]. SUS was developed as part of the usability engineering program in integrated office systems development at Digital Equipment Co. Ltd., Reading, United Kingdom. Oracle typically uses the SUS questionnaire as part of its usability methodology on software going through the usability labs in a post-release form.

3.3 SUMI vs. SUS Equivalency Scores

SUMI normative scores have been standardized against a proprietary dataset of other products. The SUMI median score of 50 represents an average score, meaning roughly half the products in the dataset have higher perceived usability and half have lower perceived usability. The product names and dates of data collection are not shared.

The SUS questionnaire is open-source with similar psychometric properties as the SUMI (valid and reliable). There is not a proprietary set of SUS products. However, because the SUS has been in use for more than 20 years, there are publicly available sets of data to generate a similar benchmark as the SUMI [6]. In addition, it has been shown that the SUS and SUMI have a strong correlation ($r > 0.8$)—a correlation that suggests that the majority (> 64 percent) of variation in SUMI scores can be accounted for by SUS scores [5].

The strong correlation between SUMI and SUS scores enabled us to generate a SUMI score that we expected to correspond closely to a SUS score. To generate our SUMI equivalency score from the SUS standardized questionnaire, we used the publicly available mean and standard deviations from [6], [7], and [8]. This data set represented more than 200 products and applications from dozens of companies and domains, rivaling the diversity and size of the SUMI database.

We took the raw SUS scores and converted them to a standardized score by subtracting each score from the mean and dividing this result by the standard deviation. The resulting normal scores (z-scores) were then converted to percentages similar to the SUMI.

To confirm the relationship between SUS and SUMI scores, we administered both questionnaires to users following a usability test and found the relationship to hold ($r > .5$), further validating the SUS equivalency scores.

3.4 Improved Usability in Oracle EBS 12

Much of our data came from Oracle EBS 11 releases 11.5.9 and 11.5.10. There has been substantial improvement to many features and functions in Oracle EBS 12, as well as focused improvement on usability. To understand how much usability improved, we gathered data on both the older (11.5.10) and newer releases (12) of two applications. We found both SUMI and SUS scores to be significantly better for Oracle EBS 12 ($p < .05$), with an approximate improvement in usability of 10 percent (the average improvements seen between SUS and SUMI questionnaires).

3.5 Sample Size

While a larger sample size is better, it has been shown that sample sizes of around 10 provide a tolerable amount of variability in the estimate of the central tendency of the scores. The number 10 is not a hard cut-off; rather, it is an approximate point in the same way that the actual variability in the mean score is a function of both the sample size and variability between respondents scores [9]. Therefore, sample sizes of 10 +/- 2 will likely not have a material difference in the stability of the mean.

We collected SUS and SUMI survey results on the customer's existing Oracle EBS 11 release 11.5.10 applications, as well as SUS and SUMI survey results for Oracle EBS 12 applications. The results gave us user feedback on Oracle EBS 11 release 11.5.10 and Oracle EBS 12 applications. We collected both SUMI and SUS data to confirm the correlation between the two scores. After all 10 sessions were completed, we analyzed the data and provided the necessary findings report on the two applications to be included in the RFT. In addition, we provided SUMI equivalency scores for the SUS data that we had for the other seven applications. To generate the SUMI equivalency scores from the SUS data, we used the publicly available mean and standard deviations from [6], [7], and [8].

4 Key Success Factors

4.1 Cross-Functional Team

Assembling a cross-functional team of a sales consultant, technical support engineer, usability engineer, and statistical analyst was crucial to the success of this project. Working together, we were able to quickly create and validate the materials needed, set up the test environment, run the study, and provide the data needed by the RFT deadline. Here are descriptions of each team member's responsibilities:

- **Sales consultant.** The project started with the sales consultant identifying usability as a differentiating factor in the prospective sale and reaching out to Oracle's Applications User Experience team for assistance. The sales consultant, located in the same country as the prospective customer, served a number of functions. She provided the rest of the team with the history between Oracle and the prospective customer, identified the specific type of data needed for the RFT, reached out to the existing customer with users matching the user profile, helped seed the data into the test environment, and served as a document translator.
- **Technical support engineer.** The technical support engineer set up the test environment in the users' native language, helped seed the sample data, and ensured the environment's stability for the duration of the usability study.
- **Usability engineer.** The usability engineer identified the necessary user profile appropriate for the two applications that lacked SUMI data and obtained the sales

consultant's validation. The usability engineer also created the task list, test plan, screener, and introductory script. Once on site at the customer location, the usability engineer moderated the 10 sessions, collected the SUMI and SUS scores, and sent the data to the statistical analyst.

- **Statistical analyst.** Before the usability test, the statistical analyst assisted in acquiring the SUMI and SUS collection documents in the users' native language. After the data was collected, the statistical analyst provided the data in the format required by the RFT and included analysis to support the correlation between SUS and SUMI data.

4.2 User Experience Group as a Valid Partner to Sales

Another key factor was demonstrating to the sales team and the prospective customer that the Oracle Applications User Experience team possessed the agility and expertise to meet its needs in the 2-week time frame provided. By providing timely, real-world support for an important prospective sale, we dispelled any misconceptions that usability work is slow, inflexible, or applicable only in a lab.

In addition, by engaging with the existing customer while collecting the needed usability data on site, the usability engineer fostered a lasting relationship with the customer and additional contacts in that region, resulting in additional activities for the user experience team, such as usability test sessions organized for the region's Oracle user group conference.

4.3 Usability Data as a Sales Differentiator

Although a number of other dimensions were included in the RFT, we later discovered that the key competitor in this sale was not able to provide any usability data for its enterprise applications. For the usability section of the RFT, the competitor received no points, thus significantly impacting its overall score on the RFT. Oracle ended up winning the contract.

5 Conclusions

As countries continue to adopt regulations that require companies to provide a workplace environment that complies with ISO standards, usability metrics will increasingly become a factor for companies in choosing which enterprise software solution to adopt. Enterprise software companies will benefit from (1) making it standard practice to use commonly accepted methods like administering the SUMI and SUS questionnaires to collect data on their key products and (2) being prepared to quickly assemble agile teams to collect additional data should the need arise.

References

1. Kirakowski, J., Corbett, M.: SUMI: the Software Usability Measurement Inventory. Brit. J. Ed. Technol. 23, 101–102 (1994)
2. International Organization for Standardization: ISO 9126-2. Software engineering — Product quality — Part 2: Internal Metrics (2005)

3. International Organization for Standardization: ISO 9241-110. Dialogue principles: Guidance on usability (2006)
4. Brooke, J.: SUS: a quick and dirty usability scale. In: Jordan, P., Thomas, B., Weerdmeester, B., McClelland, I. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor and Francis, London (1996)
5. Hoyler, A.: Methods for Evaluating User Interfaces. Doctoral Research, University of Sussex, Brighton (1993)
6. Bangor, A., Kortum, P., Miller, J.: An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 24, 574–594 (2008)
7. Lewis, J., Sauro, J.: The Factor Structure of the System Usability Scale. In: Human Computer Interaction International Conference, pp. 94–103. Springer, Heidelberg (2009)
8. Tullis, T., Albert, B.: Measuring the User Experience: Collecting, Analyzing and Presenting Usability Metrics. Morgan Kauffman, Burlington (2008)
9. Tullis, T., Stetson, J.: Comparison of Questionnaires for Assessing Website Usability. In: Usability Professionals Association (UPA), Minneapolis, vol. 2004, pp. 7–11 (2004)

ISO 25062 Usability Test Planning for a Large Enterprise Applications Suite

Sean Rice, Jatin Thaker, and Anna M. Wichansky

Oracle USA, 500 Oracle Parkway MS3op3, Redwood Shores, California 94065 USA
Anna.Wichansky@oracle.com

Abstract. In setting out to perform summative usability testing on a new suite of more than 100 enterprise software applications for 400 different user roles, we faced a number of challenges in terms of staffing, scheduling, and resources. ISO 25062 provided a valuable organizing framework to plan, scope, and implement our testing effort. In this paper, we discuss the considerations and steps that we took in test planning and management, including our prioritization strategy and creation of an automated data collection system to minimize impact on staffing resources and the usability engineering workload.

Keywords: ISO 25062, Common Industry Format, summative usability testing, enterprise software, user experience, test automation.

1 Introduction

1.1 Oracle's Next-Generation Business Applications

Oracle is a key provider of software applications that help run enterprises. Enterprise applications are typically used by large organizations, such as industries, governments, and universities, with many employees performing a wide variety of functions. Enterprise software is developed as integrated program suites that enable organizations to perform financial and manufacturing operations, control governance, and manage internal functions, such as human resources. Enterprise software also enables organizations to interface with their customers and with other organizations for procurement, supply chain management, marketing, sales, and other purposes.

Oracle Fusion Applications is a suite of industry-standard enterprise software applications, providing functionality for many different enterprises and geographies. The key features of the applications in this suite include service-oriented architecture, a role-based user experience, and embedded business intelligence. Oracle Fusion Applications is intended to replace older generations of multiplatform applications currently supported by Oracle. The Oracle Fusion Applications design captures the best user experience features of Oracle's current application suites, such as Oracle E-Business Suite, PeopleSoft, Siebel, and JD Edwards. In addition, Oracle Fusion Applications incorporates the latest in technological features, such as social networking, tagging, and superior levels of product integration to create a seamless, productive, and enjoyable experience for enterprise users.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 185–192, 2011.
© Springer-Verlag Berlin Heidelberg 2011

The Oracle Fusion Applications user experience was five years in the making. The development of this suite included an extensive and comprehensive user experience design process: ethnographic research, low-fidelity workflow prototyping, high-fidelity user interface (UI) prototyping, iterative formative usability testing, development feedback and iteration, sales and customer evaluation throughout the design cycle, and summative usability testing. This paper describes how the summative testing was scoped and planned using ISO 25062, Common Industry Format (CIF) for usability test reports [1] as an organizational framework.

1.2 What Is a CIF Usability Test?

CIF refers to the internationally standardized method for reporting usability test findings used by the software industry. The CIF is based on a formal, lab-based test that is used to benchmark the usability of a product in terms of human performance and subjective data. The CIF was developed and is endorsed by more than 375 software customer and vendor organizations led by the National Institute for Standards and Technology (NIST), a US government entity. NIST sponsored the CIF through the American National Standards Institute (ANSI) and International Organization for Standardization (ISO) standards-making processes. Oracle played a key role in developing the CIF.

The CIF report format and metrics are consistent with the ISO 9241-11 [2] definition of usability: "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." Our goal in conducting CIF tests is to measure performance and satisfaction of a representative sample of users on a set of core tasks and to help predict how usable a product will be with the larger population of customers.

1.3 Why Perform CIF Testing?

The overarching purpose of the CIF for usability test reports is to promote incorporation of usability as part of the procurement decision-making process for interactive products. Examples of such decisions include purchasing, upgrading, and automating. CIF provides a common format for vendors to report the methods and results of usability tests to customer organizations. CIF provides a benchmark of usability for the product and enables customers to compare the usability of our software to that of other suppliers. CIF also enables us to compare our current software with previous versions of our software. Our biggest reason for conducting CIF studies is that our customers expect Oracle to provide CIF reports for our products. CIF is often required as part of a Request for Proposal (RFP), especially in Europe.

2 How to Apply CIF Methods to a Large Product Suite

2.1 Project Scoping

Oracle Fusion Applications comprises more than 100 modules in seven different product families. These modules encompass more than 400 task flows and 400 user roles. Due to resource constraints, we cannot perform comprehensive CIF testing across the entire product suite. Therefore, we had to develop meaningful inclusion criteria and work with other stakeholders across the applications development organization to prioritize product areas for testing. Ultimately, we want to test the product areas for which customers might be most interested in seeing CIF data. We also want to build credibility with customers; we need to be able to make the case to current and prospective customers that the product areas tested are representative of the product suite as a whole.

Our first goal was to identify what a reasonable minimum would be for the number of CIF tests to conduct. Through discussions with internal stakeholders, including people who were most familiar with requests for CIF data that Oracle had received from customers or sales, we proposed CIF testing for 50 percent of the products in the application suite. It was understood that testing would cover only the top use cases for each product. Tests would not be comprehensive across all areas of the products chosen.

The primary activity in the scoping process was to work with the individual product teams to identify the key products and business process task flows in each product to test. We prioritized these products and flows through a series of negotiations among the user experience managers, product strategy, and product management directors for each of the primary product families within the Oracle Fusion Applications suite (Human Capital Management, Supply Chain Management, Customer Relationship Management, Financials, Projects, Procurement, and Setup).

Considerations for CIF study prioritization included:

- The products identified as most important by the strategy directors from a revenue perspective (that is, the company's strategic financial goals)
- Task flows that the strategy team identified as most important within these products
- Level of involvement of the Oracle Fusion Applications User Experience team in UI design (the User Experience team has limited resources and does not have the capacity to provide UI design for every part of the product suite)
- Previous benchmark testing, including CIF tests for previous versions of Oracle E-Business Suite and Total Ownership Experience (TOE) tests for PeopleSoft applications (we wanted to compare the usability of the new products with the previous versions where we had the data)
- Areas for which formative usability studies were conducted during the design cycle (we wanted to collect performance measures for the product areas where we believed that we could improve usability the most)
- The primary user roles for the task flow (if the product task flow was to be used by two user groups, we counted that as two CIF studies for planning purposes due to the increase in the number of participants needed to complete the usability test)

The end result of the scoping exercise was a list of 47 proposed CIF tests for the product suite.

2.2 Resource Considerations

Since 1994, Oracle has established 20 usability labs in nine locations worldwide. Each lab is supported by a dedicated lab administration and participant recruitment team and is available to all Oracle usability engineers and design teams. Because usability testing of previous applications suites was supported at these labs, it was a given that we had appropriate facilities available within the company to perform testing of the scope required for Oracle Fusion Applications.

Fig. 1. Usability engineers in darkened control room at one of Oracle's usability labs in Denver, Colorado, observe testing through a one-way mirror

Fig. 2. A test participant completes tasks during a usability test on the lighted side of a one-way mirror in one of Oracle's usability labs

We explored several options for the CIF testing, each with a different impact on time, budget, and other planned user research activities. The staffing options included using existing Oracle usability engineers to conduct the CIF testing, hiring additional full-time staff, hiring a group of contractors, or pairing interns with usability engineer leads from each product area. Each resourcing option had its pros and cons.

The additional full-time hires option was proposed to management as the ideal option, and the use of interns to test participants was presented as an alternative. Using interns would require a simplified and automated method of data collection and reporting, due to their relative lack of experience with CIF methods and the large number of tests required. Using contractors to conduct the studies was dropped from consideration due to concerns over budget, ramp-up time, and training issues. For all options, we concluded that we would need to hire an additional database administrator to support the product instances (see 2.4 Technical Challenges).

Remote testing (via web and teleconference) was considered as a means to collect participant data. Remote testing can reduce costs and allow for more diverse participant profiling, but can affect product performance due to network lag. There can also be small differences in the way the screens appear via web conference. For instance, we found that the blue color used to indicate table row selection was appearing as a much lighter yellow over web conference.

2.3 Test Protocol Issues

ISO 25062 provides useful guidance on how to standardize many details of the test protocol. Not only does this standardization add a higher degree of experimental control across the wide range of tests, but also it results in improved productivity and economy of test execution.

Here are some areas where we needed to develop policies and procedures not covered in detail by ISO 25062:

- <u>The role of training</u>. Enterprise applications contain complex functionality, and we did not want to measure user performance on a learning curve. Therefore, we had to develop a plan for standardizing the pretest use of in-product user assistance resources, such as help and product overview videos. We also wanted to standardize the familiarization instructions and product orientation time period allowed across all tests.
- <u>The appropriateness of providing assistance</u>. Providing any assistance is strongly discouraged in summative usability testing, due to its effects on the time and completion rate metrics. However, because test participants do get stuck from time to time and are not able to progress to the next step of a task, we adopted our protocol from earlier applications testing and allow two assists. If a third assist is requested, failure and a 0 percent completion rate on that task is noted.
- <u>The use of an expert's benchmark time for comparison</u>. In order to establish a rule for timing out test participants who progress too slowly through a task, in previous testing protocols, we have used expert users' time benchmarks. This expert is often a product manager or other Oracle staff member familiar with the task and software. We generally allow participants three times the benchmark time to complete a task. Thereafter, the user is timed out (and a 0 percent completion rate for that task is noted).

2.4 Technical Challenges

In order to test the Oracle Fusion Applications user experience, an entire integrated suite of related programs from a product family, such as financials, must be installed and configured. Setting up an enterprise application suite requires several pieces of software that work together to provide the suite. This setup requires a technology stack that includes: (1) a database of enterprise data; (2) the middleware, which performs most of the functional calculations and provides the services required for enterprise operations; and (3) the applications being tested. Oracle Fusion Applications runs in a web browser, and customers may use any browser certified by Oracle to work with the applications. The entire integrated suite of applications running on top of the database and middleware layers is known as an instance or environment.

Often the biggest challenge is building an instance that resembles the real world environment at a customer location. This environment must remain stable during the course of the test, and every participant must be presented with an identical environment.

From our previous applications suite usability tests, we determined two approaches to resolve these challenges:

- Setting up a system out of the box (that is, setting up a system from installation media and configuring the system post-installation with appropriate updates and additional set up to mimic the real world environment)
- Creating a clone of a real world environment on the usability labs systems

Both of these approaches require the skills of a database administrator (DBA) to create the applications environment. Creating an applications environment includes setting up the database tier with the appropriate version of the database and ensuring that the requisite database objects and database patches are installed. Following this, the DBA must install and configure the middleware components that are required to service the applications that get deployed into the middleware. The final step in setting up the environment is to deploy the applications that are being tested, and perform any profile tasks that are required to set up the environment to mimic a real world scenario. Profile tasks include setting up users and privileges.

Prior to conducting any testing, the DBA must also ensure that a backup of the database is captured and is available to be used to reset the environment prior to testing each participant. This backup ensures availability of the same exact test environment for all participants.

2.5 Automation to Facilitate Testing

The Oracle usability labs have developed and supported online data collection technologies since 1999. In order to facilitate the large number of tests required for Oracle Fusion Applications, we created a prototype online data collection system using the Oracle database and Oracle Application Express, a tool and development platform to build lightweight applications. Oracle Application Express is shipped along with the Oracle database.

This prototype system interfaces the existing usability lab scheduler with a new usability lab data logger that was built from scratch to automatically produce the CIF test report. The lab scheduler contains information about the logistics of the tests, namely what lab these tests are running in, the characteristics of the lab equipment, the number and type of test participants, and so on. The data logger contains the human performance and subjective data collected during the CIF test. Participants use an internal website to access a licensed version of the Software Usability Measurement Inventory (SUMI) [3]. We are able to capture the necessary parts of the output of these systems on a per test basis and to populate the data into an automated version of the CIF report that contains boilerplate text on the methods and available variable fields for individual and summary data.

All the underlying data is stored in an Oracle database. The information is passed to the applications like the data logger and the lab scheduler on an as-needed basis. This lab scheduler enables the usability engineer to quickly find the appropriate set up required for the usability test, and to look up participant information suited for the study.

The data logger, on the other hand, is used during the course of a study to record task and time information, as well as to record key information about the study. The data logger enables the usability engineer to quickly gather information and use the information recorded during the study to generate an automated CIF report. The automation is handled by integrating the information from the data logger and the lab scheduler.

By building an automated system using Oracle technologies, we can present analysis and metrics-based summaries of the CIF tests clearly and concisely without the usability engineer having to review every segment of the data log and to reinvent report text that is standardized across our labs. The goal is to enhance the productivity of the usability engineers and interns and provide product development teams with the findings and data analysis much sooner than traditional data collection and report creation techniques.

3 Progress and Projections

The automated CIF testing and reporting system has been successfully pilot tested in the Oracle usability labs. Formal CIF testing is expected to be underway for Oracle Fusion Applications before 2012. We project that we will be able to run the 47 CIF tests with a core group of experienced usability engineers supervising interns on up to four Oracle Fusion Applications environments simultaneously. More details will be provided at the conference presentation.

We intend to write a follow-on paper describing the results of this project in terms of the effectiveness of the automation that we have developed for the CIF testing.

4 Conclusions

Applying the CIF standards to summative usability testing for a large suite of software applications that encompass many hundred user roles and task flows is a challenging usability engineering task. However, economies can be achieved by prioritizing frequently used flows and common user roles to reduce the sheer number of tests required and by automating the testing process itself so that tests can be conducted faster with fewer and less experienced staff members.

References

1. International Organization for Standardization: ISO/IEC 25062:2006(E). Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Common Industry Format (CIF) for usability test reports (2006)
2. International Organization for Standardization: ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability (1998)
3. Kirakowski, J., Corbett, M.: SUMI: the Software Usability Measurement Inventory. Brit. J. Ed. Technol. 23, 210–212 (1994)

User Perception of Touch Screen Latency

Glen Anderson, Rina Doherty, and Subhashini Ganapathy

Intel Corporation
{glen.j.anderson,rina.a.doherty,subhashini.ganapathy}@intel.com

Abstract. The goal of this study was to determine the level at which touch screen latency becomes annoying for common tablet tasks. Two types of touch screen latency were manipulated for three applications: Web page browsing, photo viewing, and ebook reading. Initial latency conditions involved an initial delay in the screen's visual response to touch inputs but with no delay after the beginning of a touch input. Continuous latency involved continuous delay for the duration of a touch input. Both types were tested from 80 to 780 ms. Touch inputs included resizing with multitouch input, panning, scrolling, zooming, and page turning. Results showed a statistically significant main effect for application, but differences were small. Continuous and initial latency showed little difference in ratings except with ebook reading. Trend graphs show levels of user ratings by latency duration.

Keywords: touch screen latency, performance perception, multitouch.

1 Introduction

The advent of the Apple iPad™ and other tablet computers has raised interest in touch screen experience on medium sized screens. Latency in visual feedback to touch input can harm the user experience, but the current literature provides little on end-user requirements for touch screen latency. Foundational research on user perception of computing responsiveness provides basic guidelines on user performance under various conditions [1, 2], and other research has provided guidelines on sub-second task events [3, 4, 5], but it has not been specific to touch screens.

In a zero latency condition, a touch screen is a zero-order control, since it involves simple position tracking. Under latency conditions, a touch screen user must anticipate future response of inputs, thus making it a higher-order control. MacKenzie and Ware [6] utilized a target acquisition task with mouse input to show that the effects of lag and task difficulty are multiplicative in a tracking task. At their highest latency of 225 ms, performance was much lower for the most difficult target acquisition tasks. "Subjects' natural tendency to anticipate motions was severely compromised by the lag." Jay, et al. [7] utilized a tracking task to measure the effects of lag in a collaborative environment, also showing the negative impact of latency on tracking.

Touch screen use, especially for multitouch screens, is different from mouse tracking. Users make a variety of gestures on the surface of the screen, such as pinching to zoom and swipes to turn pages, which have no direct implementation on mice. The current study examined how latency impacted user ratings of system performance while doing common touch screen gestures.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 195–202, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Research Approach

Participants. Participants were 13 females and 10 males with a mean age of 32 years. All had experience with touch screens on phones or gaming systems, but none were tablet computer (e.g., the Apple iPad™) owners.

Apparatus. A 559 mm (22 in.) diagonal 3M™ Multi-Touch Display M2256PW was used to present applications and receive touch input. A portion of the screen was used to show a tablet-sized display area. The display was tilted toward the user at approximately 25 degrees to allow comfortable viewing and touch input. The applications appeared in portrait mode in an area of 168 x 224 mm, 279 mm diagonal, which was slightly larger than an Apple iPad™. The usable content display area was 265 mm diagonal for the Web Browser and 273 mm diagonal for the Photo Viewer and the E-reader. The diagonal measure was slightly less for the Web Browser due to the navigation bar at the top of the application. Due to inherent latencies in the apparatus, a base level of approximately 80 ms was the minimum touch screen latency.

A PC with an Intel Core™ 2 Extreme CPU X9650 and Microsoft Windows™ 7 drove the applications that appeared on the touch screen. The latency conditions were created by delaying the rendering of frames to the screen. The frames were buffered for a specified period of time, depending on the latency condition.

Latency. The experiment manipulated two types of latency. Initial Latency was the time from when a finger first moved until the corresponding event began to occur on the screen. If a participant maintained contact with the screen, there was no latency after the initial onset, other than the base level latency of 80 ms. Continuous Latency was also the time from when a finger first moved until the corresponding screen event, but it occurred throughout the movement of the finger.

These two types of latency correspond to issues in system architecture. When an operating system, or software running under an operating system, receives input from a touch screen programming interface, the system must determine what gesture the user most likely desired. A user's initial contact with a touch screen may not indicate the intended gesture immediately because enough input must be received to distinguish the gesture from the library of gestures that the system has. The software algorithms that interpret this input need processing time, and then the operating system requires additional time to make corresponding changes on the screen. Depending on the hardware capabilities of the system and on software design, the system will have some latency. Initial Latency reflects the "decision" time the system needs to determine what gesture is intended. Continuous Latency reflects the ongoing requirements for processing inputs through the hardware and software.

Durations for both types of latency were 180 ms, 280 ms, 380 ms, 580 ms, and 780 ms. There was 1 condition of the base level latency of 80 ms, which was Continuous, making a total of 11 latency conditions. In all latency conditions the screen updated smoothly in response to touch inputs. The frame rate, even if delayed, was high enough to avoid choppiness in the changing images on the screen.

Initial Latency was implemented by delaying the first screen update to a touch input. For example, under a condition of 180 ms latency the screen did not begin updating to match the location of the finger until 180 ms after the finger began moving on the screen in a new gesture. Once the screen started updating, it would continue to update with the gesture to the minimum level of Continuous Latency of 80 ms. The apparatus did not induce Initial Latency a second time as long as a finger remained in contact with the screen. Thus, for example, if a participant changed direction of an input, a new Initial Latency was not introduced. Thus the implementation of Initial Latency for this study may have been less annoying than a system that has additional Initial Latency for every change in gesture.

Continuous Latency was implemented by continuously delaying updates to the screen. Each frame to be drawn under the operating system was buffered for the given latency condition. Thus at any point in a touch gesture the screen would be that much time behind the finger.

Applications. Participants performed tasks with three applications: Web Browser, Photo Viewer, and E-reader. Content varied (5 Web pages, 11 photo albums, 11 E-reader books) and was counter-balanced across conditions to minimize learning effects. Figure 1a shows a representation of the Web Browser interface with a typical page loaded. HTML pages from actual newspapers were used and do not appear here due to copyright issues. Text appeared 2.4 mm high when the user interface launched and would get larger as the participant zoomed in. The tasks required only touch inputs on a single Web page – there were no live HTML links to click.

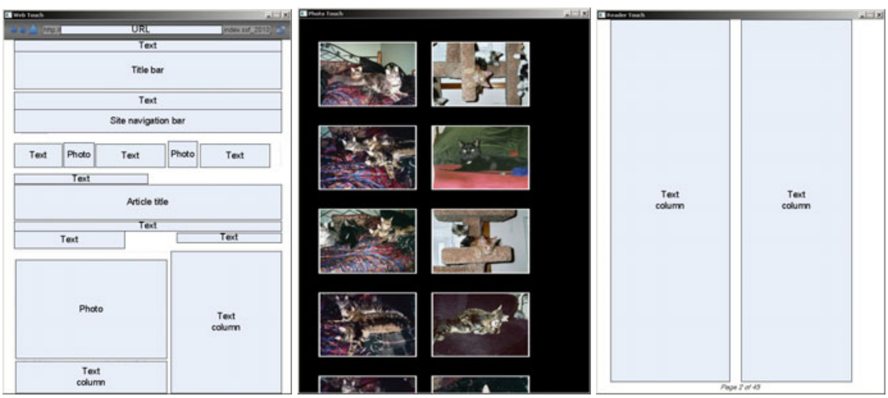


Fig. 1. 1a, 1b, 1c. Application user interfaces for Web Browser, Photo Viewer, and E-reader, respectively.

In the Web Browser application the facilitator instructed participants to use multiple touch gestures to control the screen:

- Finger and thumb in a multitouch pinch gesture to zoom in and out on the Web page.
- Single-touch movements to pan within the page.
- Single-touch swipes to scroll up and down.
- The participant could optionally hold a pinch after a resize and move the finger and thumb together to pan within the page.

The instructions were as follows:

- Zoom in and move the left column of the article so that it fills the whole browsing area.
- Scroll down to the bottom of the page.
- Scroll up and find the “Related Comments” (or “Related Articles”) portion in the middle of the page.
- Zoom out and move so that you can see the whole page again.
- Scroll up to the top of the page.
- Zoom in and center one of the images near the top of the article or in the header area side. Enlarge that image to fill the width of the browser.

Figure 1b shows the starting point for the Photo Viewer. After touching a photo, that photo would be the only item appearing in the Viewer. Participants were instructed to use touch gestures in a similar way as the Web Browser:

- Finger and thumb in a multitouch pinch gesture to zoom in and out of the photo.
- Single touch movements to pan within the photo.
- A locked pinch gesture to pan within the photo.

The instructions were as follows:

- In the photo viewer, touch the first photo.
- Zoom in on one portion of the photo, then move that portion so it is in the middle of the screen.
- Zoom out and until you can again see all photos.
- Select a second photo.
- Zoom in on one portion of the photo, then move that portion so it is in the middle of the screen.
- Zoom out so you can again see all photos.

Figure 1c shows the E-reader. Text appeared 4 mm high. Participants were instructed to use vertical or horizontal swipe gestures to turn pages. Swiping with a hold at the end of the gesture caused pages to turn quickly under all latency conditions. This approach was suitable for turning pages quickly but did not allow enough control to select individual pages. The instructions were as follows:

- Use your finger to turn the pages of the book. Turn several pages to go further in to the book and to come back to the beginning.
- Page ahead again to look for the bold text at the beginning of the first chapter.
- Now go further ahead to look for the bold text at the beginning of the second chapter.
- Turn pages to go to page 4.
- Turn pages to go to page 20.

Experimental Design. Participants worked with the facilitator individually and began sessions with practice. The goal of the practice was to allow participants to become knowledgeable and comfortable with using the touch screen. The facilitator coached participants until they were able to perform all gestures without assistance.

The experimental design was fully repeated measures – all participants experienced all 11 latency conditions under all 3 applications for a total of 33 conditions, counterbalanced to reduce learning effects.

The dependent variables were the participant ratings of each Gesture Type for a given trial, using the descriptions in Table 1. Participants rated Gesture Types as appropriate to the application (see Table 2). For example, after experiencing a latency condition with the Photo Viewer, they would rate the system for Pan/move usability, Zoom usability, and Overall Usability. Since the E-reader did not require Pan/move or Zoom, those Gesture Types were not rated for E-reader conditions.

Table 1. Rating scale

Rating	Description	Extended Description
5	Excellent	Usable with no annoyances
4	Good	Usable with minor annoyances
3	Fair	Usable but with multiple annoyances
2	Poor	Barely usable with many annoyances
1	Bad	Unusable – would not continue

Table 2. Sub-tasks used in applications

Gesture Type	Web Browser	Photo Viewer	eReader
Scroll	yes	no	no
Pan/move	yes	yes	no
Zoom	yes	yes	no
Turn Pages	no	no	yes
Overall Usability	yes	yes	yes

3 Results

Web Browser and Photo Viewer Analysis. To allow a balanced model for this analysis, Web Scroll was left out (Scroll was not rated in the Photo Viewer application), and E-reader data was analyzed separately. Web Scroll followed Web Zoom very closely, so offers little additional information. A 2 (Application) by 2 (Latency Type) by 5 (Latency Duration) by 3 (Gesture Type) repeated measures analysis of variance showed statistically significant main effects for Latency Duration ($F(4, 8)=27.10$, $p<.01$)) and Gesture Type ($F(2, 44)=2.16$, $p<.05$)). The Gesture Type by Latency Duration interaction was also statistically significant ($F(8, 176)=2.24$, $p<.05$).

Figure 2 shows the mean ratings for each Latency Duration by Gesture Type. At 580 ms of Latency Duration the Pan/move gesture was rated slightly lower. The 80 ms base latency is shown in the figure though it was not in the statistical analysis.

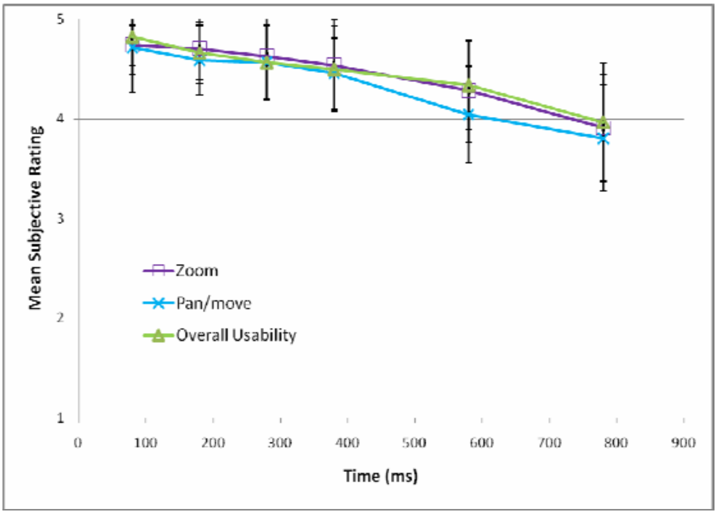


Fig. 2. Mean ratings for Zoom, Pan/move, and Overall Usability (error bars denote 95% confidence intervals)

E-reader analysis. A 2 (Latency Type) by 5 (Latency Duration) by 2 (Gesture Type) repeated measures analysis of variance showed statistically significant main effects for Latency Type (F(1, 22)=10.43, p<.01)), Latency Duration (F(4,88)=6.74, p<.01)), and Gesture Type (F(1,22)=14.31, (p<.01)). The Gesture Type by Latency Type interaction was statistically significant (F(1, 22)=4.50, p<.05)), as was the Latency Type by Latency Duration interaction (F(4,88)=2.74, p<.05)). Though the Gesture Type by Latency Duration interaction was significant, the differences among the means, relative to Gesture Type, were not practically significant. Differences between means for the main effect of Gesture Type were also not practically significant.

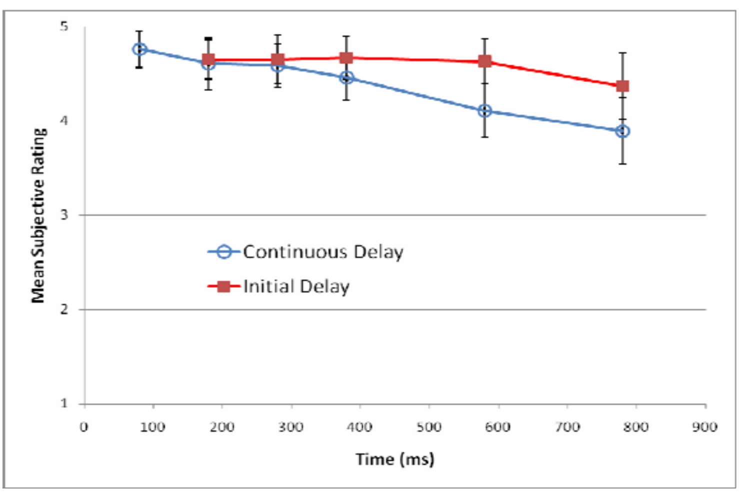


Fig. 3. Mean ratings for e-reading (error bars denote 95% confidence intervals)

Figure 3 shows the Latency Type by Latency Duration interaction. At the higher levels of Continuous Latency, the mean ratings dropped lower than under Initial Latency. The 80 ms base latency is shown in the figure though it was not in the statistical analysis.

Ratings for Web Gesture Types. Figure 4 shows the percentage of participants who gave a 3 rating or lower for Web Browser conditions, which had some of the lowest mean ratings in the study. Though the mean ratings in Figure 2 stayed at or above 4 for 580 ms Latency Duration, up to 30% of participants gave ratings of 3 or lower.

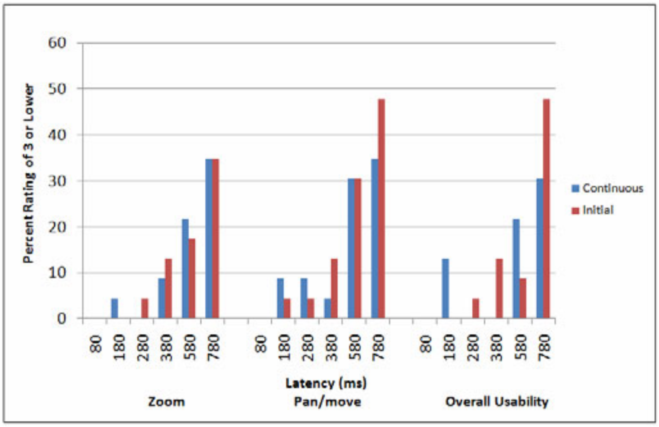


Fig. 4. Percent rating of 3 or lower for Web Browser

4 Discussion

Participants were satisfied with touch screen latency at relatively high levels. Two factors could help explain this. One, under all levels of latency, the screen used in this study updated smoothly. In other words, the rate of delivery of frames displayed on the screen was high enough so that it did not look choppy. Two, even though all participants had had some level of experience with small touch screens, most had no experience with tablets. The experience of using a tablet was novel and therefore often interesting and positive, even when the latency was high. Additional studies with users who have been exposed to tablets would show whether expectations will rise as people are exposed to these devices more.

The data showed that 580 ms as a requirement for touch screen latency may be too high for a significant portion of individual users (see Figure 4). To ensure acceptance, latency should be somewhat lower than 580 ms for a group like these participants.

In general, participants were only slightly more annoyed with Continuous Latency than with Initial Latency, even though Continuous included Initial Latency at the beginning of each new touch input. In this study Gesture Types (Zoom, Pan/move, Page Turn) were relatively brief, as is common in current real-world tablet usage. It is possible that longer gestures, like continuous tracking, would be more prone to annoyance with Continuous Latency.

E-reader page turning under Continuous Latency was an exception. The participants found these conditions less usable than Initial conditions. This may be for two reasons. First, with Initial Latency, by the time the relatively time-consuming page-turn gesture was finished, part of the latency period was over. With Continuous Latency, the system kept the latency level through the end of the page-turn gesture, thus making the user wait for the full latency period at the end of the gesture. Second, the page turning task was highly repetitive and one page turn was much like another, affording the development of more precise expectations. This may have allowed participants to make more sensitive judgments.

Since there is little published literature on touch screen responsiveness, additional research could explore several related areas:

- Screen size – Touch screens are being used in many differently sized devices.
- Other tasks – Touch screens are also used in gaming, point of sale, and other categories.
- Smoothness – Some touch screens have a slightly choppy appearance in response to touch inputs. This variable may interact with latency.
- User experience – As users are exposed to more touch screens, their standards for performance may change.
- Performance variables – Additional studies could include error rate and task completion time.

Acknowledgments. Special thanks to Daniel Theophanes for development of the test platform.

References

1. Miller, R.B.: Response time in man-computer conversational transactions. In: Proc. AFIPS Fall Joint Computer Conf., vol. 33, pp. 267–277 (1968)
2. Shneiderman, B.: Response Time and Display Rate in Human Performance with Computers. ACM Computing Surveys 16(4), 265–285 (1984)
3. Seow, S.C.: Designing and Engineering Time. In: The Psychology of Time Perception in Software. Addison-Wesley Professional, Reading (2008)
4. Dabrowski, J., Munson, E.V.: Is 100 Milliseconds Too Fast? In: Proc CHI 2001, Extended Abstracts, pp. 317–318 (2001)
5. Anderson, G.J., Doherty, R., Baugh, E.: Diminishing returns? Revisiting perception of computing performance. Publication pending for Proc. CHI (2011)
6. MacKenzie, S., Ware, C.: Latency as a Determinant of Human Performance in Interactive Systems. In: Proceedings of the ACM Conference on Human Factors in Computing Systems – INTERCHI 1993, pp. 488–493 (1993)
7. Jay, C., Glencross, M., Hubbold, R.: Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. ACM Trans. Comput.-Hum. Interact (TOCHI) 14(2) (2007)

Interface for Multi-robots Based Video Coverage

Ryad Chellali[1], Emmanuel Bernier [1,2], Khelifa Baizid[1,2], and Mohamed Zaoui[3]

[1] Italian Institute of Technology, Via Morego, 30, Genova 16163, Italy
[2] University of Genova, Genova, Italy
[3] Collège de France, France
{ryad.chellali,emmanuel.bernier}@iit.it,
Baizid.khelifa@gmail.com

Abstract. In this paper we address the problem of coverage area with multi-robot system by giving an autonomy level to single robot as well as the team. The coverage concept is based on Voronoi algorithm and it applied to extract the minimum number of views from the supervised environment. Human Robot Interface was developed to allow user to selects minimum points' locations where robots have to achieve the supervision task. Moreover, user can access at any time through the interface watching video streams feedback coming from robots and semi-autonomous robots can locally correct their paths if probability collusion, with the environment, is predicted based on new adapted Active Counter Model. However, failed robots can easily request help from user, where, this last acts overcoming the robot from its unlikely situation. The coverage strategy that we proposed combines the real robot motion (by consideration motion error and process the correction method) and the human robot interaction as well as the supervision paradigms.

Keywords: coverage, Active counter model, multi-robots system, semi-autonomous robot.

1 Introduction

When dealing with urban search and rescue operations, time is the critical parameter has to cope with: the faster one finds victims the more chance has to save lives. We are developing a multi-robots system for USAR (Urban Search And Rescue) purposes. One of the main topics we are tackling is dealing with the shared control to capture video streams about specific areas operators aim to cover. In order to reduce the workload of operators, our system supports as inputs the areas to cover. After that, the system determines the whole strategy for robots to find first the optimal position for each one. In a second stage, each robot achieves its safe movement toward the specified point derived from the previous phase.

Unfortunately, the semi-autonomous mode does not correspond to real conditions. Indeed the context and the working environments are at best partially known: after catastrophes or fires, these environments are no more conformal to plans rescuers have. Following that, operators' intervention is required to modify on the fly one or more robots' strategies in order to fulfill the requirements.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 203–210, 2011.
© Springer-Verlag Berlin Heidelberg 2011

In this contribution, we detail the components of the interface we developed allowing operators to interact intuitively with robots both to specify the areas to cover as well as to support direct robots reallocation and redistribution in order to cover given areas.

2 Related Works

Several researches works have been carried out demonstrating the coverage problem. Where, the original research works has introduced by Gase [1]. This approach is an analytical technique to optimize the spread of a mobile robot group for the coverage area in a convex environment based on the Voronoi algorithm. However, this approach may not consider the challenge of environments with obstacles.

A similar approach studied the problem of optimal placement for mobile robots team in surveillance tasks for non-convex environment implementing the Voronoi method [2], robots displacement is based on the importance area estimated by sensor information. Heterogeneous robots group method [3] as extension of location optimization concept [4]. This method takes into consideration flying and ground robots. Another extension for the non-convex environment case with unknown obstacles using the combination of Voronoi and partition and potential field is also proposed in [5].

In [6] a method was proposed based on the potential field for convex environment coverage. Each robot must have a connectivity feature of the communication range of its neighbors to ensure the algorithm process. In [7] the authors introduced a method for non-convex environment based on the gradient algorithm to locate the mobile robot observer in the environment. However, this method works for single robot case.

Considering the visibility concept, an approach using multi-robot system and assumes that the environment is completely connected is presented in [8]. The visibility problem could be related to the Art Gallery Problem. This method looks for the optimum number of guards in a non-convex environment for which, each segment of the environment is visible at least from one guard [9], [10], [11].

In [12] a coverage method based on the cognitive adaptive optimization algorithm (CAO), it is dedicated for heterogeneous robots and unknown environment having obstacles.

However, most of the cited techniques give solutions on the basis of an a priori knowledge (of the robot motion and the environment), which may not correspond to reality. To cope with local uncertainties and local errors (movement and sensing errors), we proposed a method which allows robots to modify and optimize locally their routes. Moreover, in case of robots fail or detect a bottleneck situation, they inform the operator, who then takes the control of the blocked robot(s) to reprogram on the fly a new path and its feasibility is simulate by moving virtual robots within the top-view map. For our system, we use the RRT [13] algorithm to generate the a priori trajectory of the robot to reach the area to be supervised. While the mission is in progress, the robot corrects its path if needed (presence of obstacles) based on new adapted Active Counter Model [14] that we developed. In addition to smoothing the executed trajectories, the SNAKE based approach allowed us to include in our path generator, different uncertainties and modeling errors produced by on-board sensors.

3 Multi-robots Platform

We developed a multi-robots platform [15] composed of wheeled and legged robots. Each of them is endowed with a video camera and a laser-scanner. The embedded processing/sensing capabilities allow to avoid obstacles through the proximity sensors and to generate a free path given a global map. On the other hand, all robots can communicate with the central system and with each other through the wireless networking infrastructure. Last, the robots can be fully tele-operated by operators: a joystick allows controlling movements regardless to the mobility of the considered robot. As well, the same joystick allows the control of the embedded camera pan-tilt movements.

4 Supervision and Control System

To manage the robotics platform we developed a central system composed of I/O, control and supervision tools. For specifying the task, namely the coverage area, a top view map and a simple 2D input tablet are used. Operators can draw directly the concerned area; which is handled by the task generator. This later considers the a priori map to derive the free path and the final poses for each robot. We used the RRT algorithm [13] and for poses, we use Voronoi distribution an extension of the art gallery problem heuristics. Both poses and paths are sent to robots. Each of them works then in a semi-autonomous mode while reaching the goal. Operator follows the execution of the task (robots motion) on the top view map. As well, they can access to video streams sent by each robot.

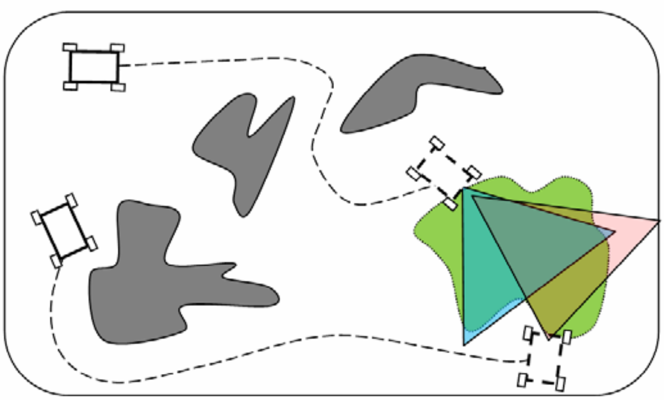


Fig. 1. An example of area coverage

4.1 Path Planning and Robots Autonomy

The path of the robots is generated using RRT algorithm, this algorithm is based on random note selection which process finally to a theoretically feasible path (without collusion). However, the generated path can work only if we consider no error on the robot motion, which is not almost the real case. So the RRT path generation method serves our platform, only, for a preliminary path definition, which could not be safe due to motion error of the real robot as well as sensor uncertainties. Considering real robot case, this last never follow theoretical path which makes the challenge to cope with this kind of problem very hard [16]. Our objective, is *first*, to unable the robot to reach a desired location without any collusion with the environment, where, the robot can modify its own path. if it predicts a collusion probability with the environment. *Second*, optimize the trajectory of the robot by minimizing the RRT path.

We proposed a new concept based on the Active Counter Model [14]. Unlike previous research used the Snake for Image Processing [17] we adapt this concept to cope with the path generation and correction. Begin; based on the force given by the obstacle the Snake modifies the path to be much safe. We build the energy of the Snake model based on the obstacles and the robot path; both are combined within a defined process. Where, the robot motion errors and the sensor uncertainties are merged in which the result led to a safe trajectory, by modifying this last.

Based on the energy minimization our Snake model will behave under the influence of the environment and the internal trajectory forces. Where, the external force of the environment (Obstacle and trajectory) push the counter outside zones having big collusion risk probability. The internal force keeps the smoothness of the counter as it possible. While, the optimization force reduces bends generated by the RRT. Assuming that the trajectory parameters are written by $v(s) = (x(s), y(s))$ The Snake energy model is defined as follow:

$$E_{Snake} = \int_0^1 E_{IntFrc}(v(s)) + \int_0^1 E_{ExtFrc}(v(s)) + \int_0^1 E_{OpFrc}(v(s)) \tag{1}$$

Where, E_{IntFrc} represents the internal energy of the counter, Where, E_{ExtFrc} represents the external energy of the counter (obstacle + trajectory) and E_{OpFrc} represents the energy to optimize the counter.

The internal energy is defined as follow:

$$E_{IntFrc} = \left(\alpha(s)|v(s)|^2 + \beta(s)|v(s)|^2\right)/2 \tag{2}$$

Where the first part (function of α) is a membrane force of the counter, while the second part (function of β) is a force acting to counter to be thin.

We are using the term of to 1- the only the first part of the External energy force $E_{ExtFrc} = -\left(G_\sigma * \nabla^2 I\right)^2$ defined in [16] is defined as follow. In this case our external energy is defined as follow: $1 - E_{ExtFrc}$.

The Optimization energy $E_{OpFrc} = \int_0^1 \sum \vec{f}(x, y)dt$ is defined by the:

$$F = N * \vec{n}(s) \tag{3}$$

Where $\vec{n}(s)$ represent the normal unitary vector to the curve at point $v(s)$ and N weight, for a large value of N the curve converges to straight line very quickly. The normal of each point is strongly related to all previous and next points.

The obstacle's primitives are generating using distance sensor (URG laser) the obstacles are modeled by a polygons (fig. 2. (b)). While, the sensor error is modeled by a Gaussian function, which applied to the primitives to cope with the error Δt .

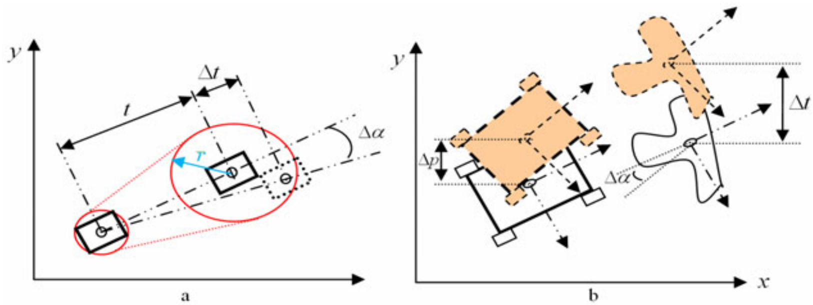


Fig. 2. Primitive concept, (a) path motion, (b) position tracking on the obstacles primitives

Trajectory primitive are represented regarding to the error motion of the robot, where, the position $(x, y)_i$ is related to the length of the executed motion from $(x, y)_{i-1}$ to $(x, y)_i$. So the robot can reach any position in the red circle. Fig. 2.(a) is showing the envelope of robot motions which can be accumulated from position *i* to *i+1*. Because of the correction strategy of the robot, that our platform allows, the errors are not accumulated in all positions, but only between the two which are successive. The radius of the red circle (fig. 2.(a)) which represents the possible reached points from the robot related to the uncertain position $(x, y)_i$ is defined as follow: $r = \max\left(\Delta t / \cos(\Delta\alpha), t * \sin(\Delta\alpha)\right)$. For the uncertainty of the distance sensor the translation Δt represented in fig. 2.(b) because of the error in the tracking system, which calculate the robot position using SLAM technique [18]. Where, the robot position uncertainty Δp affects an error on the obstacle because this last is represented in the robot scale.

4.2 Robot Placement Algorithm

The aim of this part is to find the minimal number of sensor position from which the full coverage of the targeted area is effective. The problem is similar to the classical AGP problem, where one determines the minimal number of guards and their positions to watch all gallery walls. Let us consider a camera as the used sensing system to perform the coverage. Let us as well consider that the environment is polyhedral. In this case, a classical Voronoi diagram can provide the set of points from which one is guarantied to achieve the watching task (fig.3a). Indeed, the external Voronoi vertices (fig.3b) are equally situated from objects borders and belong the respective polygons sides' bisectors. Obviously, this set is not optimal and overlaps may appear. To reduce the number of points, we developed a visibility propagation algorithm [19]. The process starts by detecting the minimal points that one object needs to be covered (parts in red in fig-3c). For the presented case, we have 4 points. These points are used to propagate the visibility to the other objects. Once a part of any polygon is seen, it is removed. We rerun the first step on a new object: for object (2) in fig-4d, we have 3 new points to fully cover the object. We propagate the visibility to the two lasting objects. In fig-3e and fig-4f we have respectively two and two new points. In total we have 12 points to cover the full area regardless to concavities, occlusions or other classical limitations.

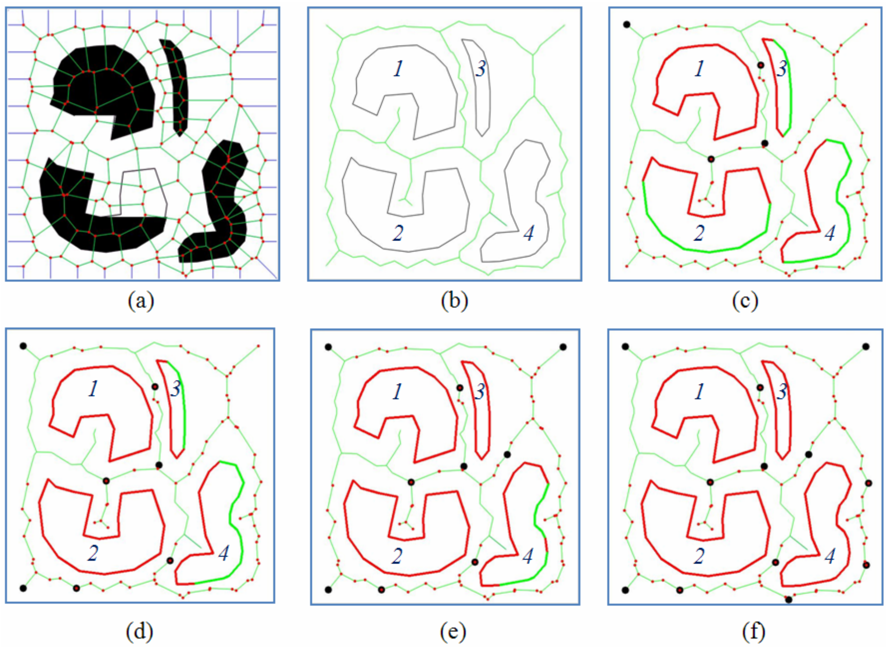


Fig. 3. Coverage area algorithm example, (a) external Voronoi vertices distribution, (b) initial points configuration, (c) to (f) the first, second, third and fourth object coverage respectively

5 Experiments Scenario

The full platform was tested in our laboratory conditions to assess and to validate all the developed blocks. Currently we are comparing between the results one can obtain through the full tele-operated mode and the semi-autonomous mode in order to show the effectiveness of the developed algorithm (fig. 4).

An example of the self-correction of the robot path is shown in fig. 4. By using the interaction interface user select the robot in (fig.4b) and then start the action in (fig.4a) by showing the target in (fig.4c). After generating the path primitive (fig.4d-right) a Gaussian process is applied on this path as well as to the obstacles. Based on this the external force will be computed (fig.3d-up right). (fig.4d-right-sown) contains the original path with black color (generated by the RRT), the Snake's optimized path with green color and the real robot path execution with the red color. In this case the Snake has pushed the path to be more save (fig.4d-right-up).

However, an unknown blue obstacle (because of the occlusion problem) makes the robot staked during the execution, where, the operator has fully tele-operates the robot (blue portion) to help it to overcame this problem, and then the robot has continues exciting the path (last read portion) until arrived to the target.

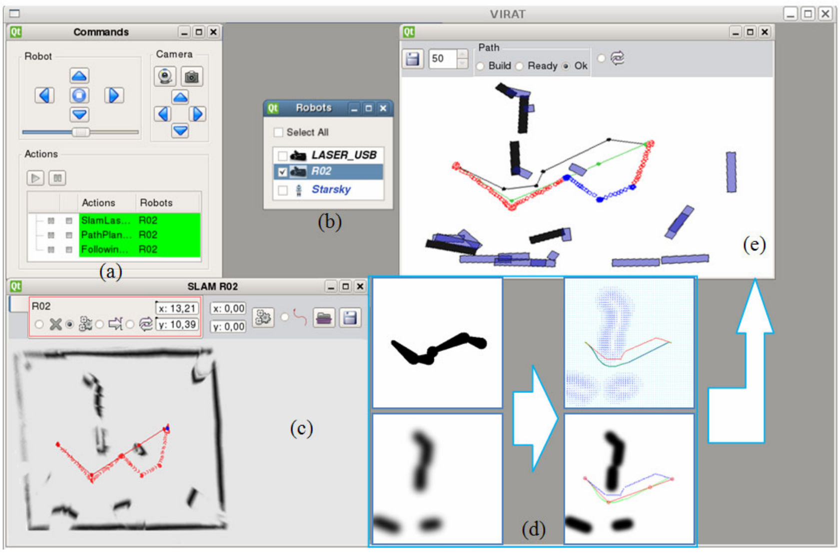


Fig. 4. Interaction interface with safe path using Snake concept, (a) command windows, (b) robots selection, (d-right) path and obstacle primitives and their Gaussian distribution, (d-left) external force, (e) original, Snake's and real paths

6 Conclusion

A new method for coverage area considering a shared and robust control between a real robot and human operators has been proposed. The goal was to optimize the task by extracting, the minimum points of views as a first level, while the second optimization level is done with the robots' path using new model of SNAKE concept. From the robot autonomy point of view, results show that the real robot motion and distance sensor uncertainties are well handled and the robot can modifies its own path to be much more safe by minimizing collisions probability. The proposed combination between of the human supervision and robot autonomy, allowed performing successfully the task, even with unknown situations that the real robot has found.

References

1. Douglas, W.: Gage: Command control for many-robot systems. In: Nineteenth Annual AUVS Technical Symposium, Hunstville Alabama, USA, pp. 22–24 (1992)
2. Schwager, M., McLurkin, J., Rus, D.: Distributed Coverage Control with Sensory Feedback for Networked Robots. In: Proceedings of Robotics: Science and Systems, Philadelphia, USA (2006)
3. Pimenta, L.C.A., Kumar, V., Mesquita, R.C., Pereira, G.A.S.: Sensing and Coverage for a Network of Heterogeneous Robots. In: Proceedings of the 47th IEEE Conference on Decision and Control, Mexico, pp. 3947–3952 (2008)
4. Cortes, J., Martınez, S., Karatas, T., Bullo, F.: Coverage Control for Mobile Sensing Networks. IEEE Transactions on Robotics and Automation 20, 243–255 (2004)
5. Renzaglia, A., Martinelli, A.: Distributed coverage control for a multi-robot team in a non-convex environment. In: IEEE IROS09 3rd Workshop on Planning, Perception and Navigation for Intelligent Vehicles, St. Louis, MO, USA, pp. 76–81 (2009)
6. Poduri, S., Sukhatme, G.S.: Constrained Coverage for Mobile Sensor Networks. In: IEEE International Conference on Robotics and Automation, New Orleans, LA, USA, pp. 165–172 (2004)
7. Ganguli, A., Cortes, J., Bullo, F.: Maximizing visibility in nonconvex polygons: nonsmooth analysis and gradient algorithm design. In: American Control Conference, Portland Oregon, USA, vol. 2, pp. 792–797 (2005)
8. Ganguli, A., Cortes, J., Bullo, F.: Visibility-based multi-agent deployment in orthogonal environments. In: American Control Conference, New York, USA, pp. 3426–3431 (2007)
9. Ferrari, S.: Track Coverage in Sensor Networks. In: American Control Conference, vol. 2, pp. 14–16 (2006)
10. Shermer, T.C.: Recent results in art galleries. IEEE Proceedings 80, 1384–1399 (1992)
11. O'Rourke, J.: Art Gallery Theorems and Algorithms. Oxford University Press, New York (1987)
12. Renzaglia, A., Doitsidis, L., Martinelli, A., Kosmatopoulos, E.B.: Cognitive-based adaptive control for cooperative multi-robot coverage. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3314–3320 (2010)
13. Lavalle, S.M., Kuffer, J.J.: Rapidly-exploring random trees: Progress and prospects. In: Workshop on the Algorithmic Foundations of Robotics (2000)
14. Kass, M., Witkin, A., Terzopoulos, D.: Snakes Active contour models. Journal of Computer Vision, 321–331 (1987)
15. Baizid, K., Zhao, L., Mollet, N., Chellali, R.: Human Multi-Robots Interaction with High Virtual Reality Abstraction Level. In: 2nd International Conference on Intelligent Robotics & Applications, Singapore (2009)
16. Bouilly, B., Simeon, T., Alami, R.: A numerical technique for planning motion strategies of a mobile robot in presence of uncertainty. In: IEEE International Conference on Robotics and Automation, vol. 2, pp. 1327–1332 (1995)
17. Aizu, M., Nakagawa, O., Takagi, M.: An image segmentation method based on a color space distribution model. In: International Conference on Mechatronics and Automation, China, pp. 532–533 (2006)
18. Steux, B., Oussama, H.: Oussama.: tinySLAM: A SLAM algorithm in less than 200 lines C-language program. In: International Conference on Control Automation Robotics Vision, pp. 1975–1979 (2010)
19. Chellali, R., baizid, K.: New Algorithm for Coverage Problem. Technical Report, IIT (February 2011)

Precise Mishandling of the Digital Image Structure

Theodore Davis

The Basel School of Design, HGK FHNW,
Vogelsangstrasse 15, 4058 Basel, Switzerland
ted@teddavis.org

Abstract. In the everyday capturing and sharing of digital images, one is rarely confronted with the notion of this media being any different than its analog counterpart. With the exception of an occasional compressed artifact, the medium remains transparent to the pictorial content it contains—that is until an error or glitch occurs, providing a brief glimpse into the inherent properties of the medium itself. However, through advancements in error detection techniques, these artifacts are seen with less and less frequency, thus further removing the message from its medium. What if the glitch was caused intentionally? What range of innovative visual possibilities might unfold as the digital image is pushed beyond its ideal state? In pursuing these questions, the term precise mishandling was coined to describe a method of carefully and thoroughly manipulating the structural code found within a digital image, meanwhile observing the reaction this had on the pictorial surface. Utilizing the most common file format for digital images, the JPEG, this paper follows the discovery of a technique for generating innovative imagery through the cross-media visualization of text.

Keywords: new media, media studies, interface design, design process, image design, digital image, jpeg file format, precise mishandling, glitch, cross-media visualization, text2image.

1 Introduction

Through the use of commercial design software, most of which offer tools analogous to hand techniques, the range of visual possibilities is, to no surprise, nearly identical to that of their analog inspirations. Further exemplified by the reoccurring reference to a *canvas*, this selective focus onto the surface of an image greatly ignores the digital code of which the medium is entirely composed. Similar to the function of a window, through its transparency, it allows one to view the outside surroundings and is rarely noticed until it becomes excessively dirty or a crack occurs. At this point, the window transfers from a transparent or unnoticed medium to an opaque one. As common file formats such as the JPEG/GIF/PNG are used to store and share digital images across the Internet, awareness of the media itself is rare, albeit for minor artifacts of compression that may occur through each exchange. To the average user, this process is seemingly mundane—that is until a glitch occurs along the path of transmission, in

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 211–220, 2011.
© Springer-Verlag Berlin Heidelberg 2011

other words generating a crack in the window. As the content is disturbed, this rarely desired error produces an unpredictable visual, potentially revealing inherent qualities within the file format itself.

Through the advances of error detection and correction techniques, such a phenomenon occurs with less and less frequency, thus further removing the message from its medium [1]. However, what if the glitch was caused intentionally, manipulating the data file directly, thus circumventing the system set in place to counterbalance it? What range of visual possibilities might unfold as the medium is pushed beyond its *ideal* state?

In the pursuit of answering these questions, the term *precise mishandling* was coined in order to fit the proposed methodology. That is, the intentional and systematic manipulation of a stable system, in order to learn how it functions and to discover potential exploits for creative use. It was through this technique that an investigation of the digital image's *underlying code* [2] began, soon leading to the development of tools capable of generating a wide range of new visual possibilities.

2 Precise Mishandling in Recent Years

The term, *precise mishandling*, can be useful in describing artistic practices in which the error or *glitch* is caused with intention for creative expression. Within the fields of music [3] and visual arts [4], the glitch offers unfettered access to the unexpected and uncontrolled within a computerized or electronic system. While occurring naturally and with great spontaneity, artists have harnessed and utilized the glitch in what could be seen as a rejection of the pixel perfection that exists in our digital representation of analog artifacts. This could be found in recent years through a popularized experimental technique dubbed by many as *datamoshing* [5] in which the key frames of a movie file are removed, forcing the computer to blend pixels at will, everytime a dynamic change in the moving action takes place.

A precedent for specific file format manipulation can also be found in Cory Arcangel's 2003 work *Data Diaries* [6]. Created in a time of minimal error correction, the work consisted of removing the entire data contents from a *Quicktime* video file format, while maintaining the header used for playback instructions. According to his explanation [7], once the original content was removed, the player defaulted to displaying the computer's daily stored RAM as though it were video data. The result was a streaming array of pixels, waves and bands, as 128 megabytes of random access memory were spontaneously transcribed. This repurposing of media could be considered a form of *cross-media visualization*, in which, one form of digital media merely acts as a vessel for the display of another.

3 Exploring the Digital Image Medium

When considering the many formats a digital image can embody, the JPEG or *.jpg*, is without a doubt the most common in use. Its prevailing application can be found as the default file type used by the majority of digital cameras as well as when sharing images across the Internet. This is likely due to the standards having been set at an early stage of digital adoption, along with a compression scheme that provides a significant reduction in file size while maintaining photographic detail. As the file type remains virtually the same since its inception in the early nineties, it became the perfect candidate for exploration. While, in the beginning of this research, a series of tests were conducted on the wide range of digital image file formats currently in use,[1] the JPEG proved to be the most susceptible to manipulation, which is likely due to an entropic encoding and decoding process.

{digital image} » {*encoder specifications*} » {compressed jpeg file} .
{compressed jpeg file} » {*decoder specifications*} » {visible image} .[2]

3.1 Surface versus Structure Relationship

A core attribute that separates the digital image from its analog counterpart is that it could be devised into two modes of representation, the image surface and structural code. The image surface refers to the pictorial quality that is rendered by image viewing software, allowing it to "participate in dialog with other cultural objects" [8]. Whereas the structural code consists of the rudimentary ones and zeros of binary code, in which nearly every digital file is composed of [9]. Before this data is displayed as a digital image, it could be said to 'participate in dialog with other *computational* objects'. In other words, the binary representation of an image file is no different than that of a text, movie or sound file.

When it comes to comprehension, the computer remains limited by its near solitary understanding of the numerical code.[3] Humans on the other hand, have the ability to interpret both, even though out of convention only the surface level is read. With the

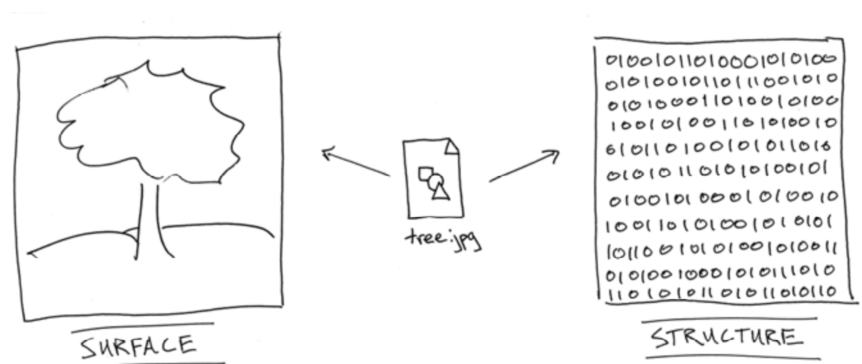


Fig. 1. The digital image file as represented by both its *surface* and *structure (underlying code)*

[1] JPEG, GIF, and PNG being the most commonly used for transferring bitmap images, while additional format tests included BMP, PICT, PDF, PSD, SGI, TIFF, and TGA.

[2] Based on *JPEG Standard recommended guidelines* of www.w3.org.

[3] While significant advances in Image Processing exist, getting beyond the lower-level attributes (color, texture, form), to a comprehensive contextual understanding by the computer, remains a distant challenge.

aid of software such as the hex-editor[4], a virtually unknown tool in the realm of visual communication, the aptitude to comprehend this structural layer is enhanced by the capacity to precisely modify the now visible code (see Fig. 1).

3.2 Comprehending the Code

Making the structural code of a digital image visible could be analogous to opening the hood on an automobile, particularly when one begins to poke and prod at the foreign matter found. Without prior experience, one is likely to break the image structure, thus rendering the file unreadable, or in the case of the automobile, immobile. However this analogy quickly dissolves as the computer, with its lack of an original in terms of duplication, can be used to learn through trial and error, while always maintaining a backup file to revert to. It was through this process that the technique of *precise mishandling* was developed. In this case, selecting every offset within the structural code and pushing its value from an initially stable state through the entire hexadecimal range (00 » FF or 256 values).[5]

Originally generating every hexadecimal variation by hand, followed by the use of a keyboard macro, and soon after a programmed script, a genetic schema of the digital image was thoroughly unraveled, resulting in a visual map of structural deviations. Parallel to reading technical descriptions about the JPEG file format [10], [11], [12], these illustrative examples of each described region now offered a better understanding from an outsider's point of view.

3.3 Dividing the Code

As is the case with all digital image files, the JPEG file format structure is divided into two distinct parts, a header and body (see Fig. 2). The *header*, which resides at

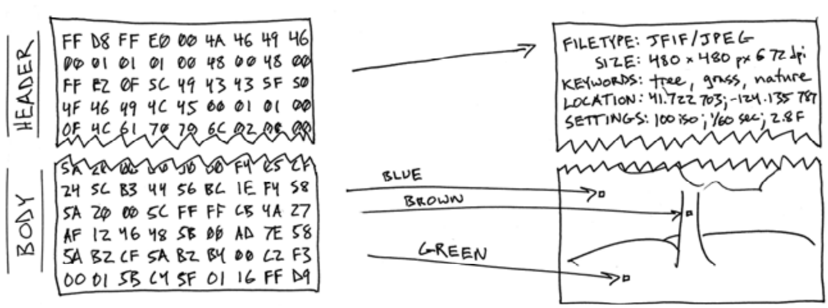


Fig. 2. The structural code is split into a *header* and a *body.* The *header* consists of rendering instructions along with metadata including declarations such as the file format used, dimensional size and resolution. The *body* is composed of literal clauses addressing specific pixels and their visual attributes.

[4] A hex-editor allows the user to view and modify the binary code (represented as both hexadecimal and ASCII character format for ease) of any data file. This software is often used in the field of digital forensics.

[5] Visit *www.ffd8.org/header_x256* to see the entire range of visual results for one particular region of JPEG file format, the *quantization table*.

the top of the image, can be described as instructional information read by any given software to decipher the type of data being processed and how it should be rendered. The *body*, which follows a designated marker signifying the completion of the header, is responsible for the actual content being processed or in this case, the actual picture. While research was conducted to both regions in parallel, this paper focuses on the findings within the *body* region.

4 Precisely Mishandling the Digital Image Body

Through learning to decipher the various regions of the JPEG structure, it became clear that the primary textual information found within the digital image, *metadata*, was being stored within the *header*, therefore sitting on the periphery of the actual pictorial contents. While this is completely logical considering the information must pass, untouched, through a decoding/encoding process, it leaves opens the potential for precisely mishandling the placement of such text. Specifically, to explore the cause and effect relationship that might occur when injecting textual content directly into the *body* of the image, prior to the decoding process. Would it be possible for the text to find its way, in recognizable form, to the image surface—or might the JPEG simply reject the content and refuse to render?

Based on the decoding process a JPEG file goes through, the hypothesis was formed, that if the *header* or instructions for rendering the image were kept intact, the system would have no choice but to display injected information in the exact same way as it would with existing data. This thesis proved correct immediately following the discovery of where the header ended and the body began.

4.1 Con*text*ual Images

With the body of the digital image now isolated, textual information could be inserted at any point within this code. Initial experiments involved the pairing of texts that had a contextual relationship to the photographs they would inhibit. This was first done with the use of presidential portraits and an accompanying transcript from the inaugural speeches given at the beginning of each term in office (see Fig. 3). Utilizing a raw unformatted text of each speech, a subjective decision was made as to where in the image this text would begin. An offset corresponding to the pictorial surface location of their mouth was used for added association. Since each Ex-President had held two terms of office, this offered two sets of sample data to inject and compare.

C_0 C_1 C_2 B_0 B_1 B_2

Fig. 3. Transcripts of the now Ex-Presidential inaugural addresses were injected into their portraits (C_0 *and* B_0) as found on whitehouse.gov in 2008. The first and second term speeches can be found in the corresponding numbered suffix, C_1 *and* C_2, respectively.

Upon the implementation of this experiment, one particular variation that occurred in both cases was the brightness of tonal value decreasing as the picture and text flow towards the bottom. As expected the hue initially begins with the same value as the mouth region in which it began, however with the first speech, this value changes both in brightness and adopts color values that were never present in the original image. This trend does not occur in the second speech, in which the tonality remains nearly the same from the mouth until the bottom of the image.

Expanding upon this study, similar experiments were conducted using an official photograph of The White House in combination with the 37 inaugural speeches given by former US Presidents. The symbolic offset of 1600 was selected[6] and each inaugural speech was injected into the same location within the image structure. Additional studies were created through the convergence of MySpace profile photos and their profile descriptions. This was done with the ambition of visualizing the parallel or disconnect in ones projected appearance both in terms of pictorial and textual quality.

With each additional experiment, so grew the impression that a contextual visualization of the text was being developed. This gave the hope that one could extract a deeper meaning from the images being produced. However, this conviction became less certain as the location of injected text was altered, even with the slightest change (see Fig. 4). In the White House experiment, changing the offset location of injected text from an initial 1600 to 6400 or further to 12800 changed the resulting image dramatically.

J_0 J_{1600} J_{6400} J_{12800}

Fig. 4. From an initial web resolution photograph of The White House (J_0) a transcript of Thomas Jefferson's first inaugural speech was injected into the digital image structure at offset 1600 (J_{1600}). This was subsequently followed by injecting the original digital image at offsets 6400 (J_{6400}) and 12800 (J_{12800}), respectively.

It was at this point that the decision to work with smaller lengths of text arose, thus isolating the visual deviation of each minute change. Individual characters were soon used to replace the entire pictorial contents within the code and thus the medium truly became the massage.

4.2 TEXT2IMAGE

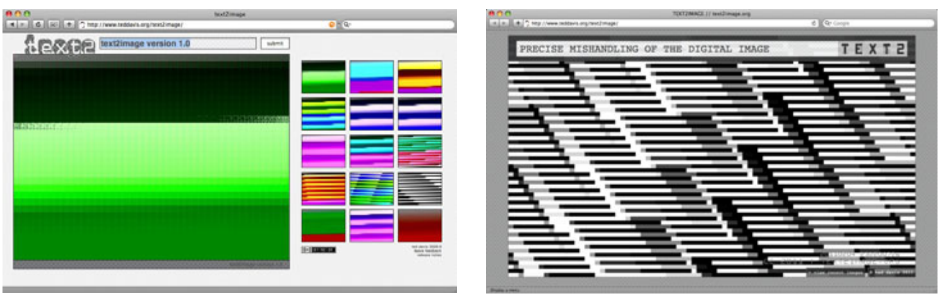
While the initial process of injecting text into the digital image was done manually, it took an absorbent amount of time to generate each new variation. Along the process of *precise mishandling*, a state of *exploration* shifted to *discovery*, at which point it became time to find an automated way to accomplish the task.

[6] 1600 Pennsylvania Ave. NW, Washington, D.C., is the physical address of The White House.

Motivation for developing an automated tool also stemmed from an interest in exploring and cataloging imagery that utilized the widest gamut of textual input possible. This included input from other languages, keyboard characters beyond the norm, and all content that would have never been conceived by one person alone.

Implemented as a web application using a LAMP (Linux, Apache, mySQL, PHP) setup, *TEXT2IMAGE* [13], a globally accessible and platform-free tool, was born. To great surprise, within the first five days of being launched, it amassed more than 13,000 images—20,000 by the second week. Designed with simplicity and universal access in mind, the interface consists of little more than an empty input textbox and a large colorful image filling the page (see Fig. 5). Upon loading the website, the textbox is automatically focused upon, ready to accept virtually any textual input from the user's keyboard by utilizing the UTF-8 character set. The enter key or submit button is pressed and within seconds the main image refreshes, displaying a new image, generated entirely from the provided text. Modifying the text ever so slightly has the potential to shift the image in dramatic ways, however the rendered image is always consistent for a given text.

Extended features of *TEXT2IMAGE* include a *recent images* tab, showcasing an array of the last fifteen generated images. Hovering the mouse over each thumbnail displays the text that was used to create it, whereas clicking will replace both the main image and textbox with the selected output. This allows one to continue exploring from a point of existing visual inspiration. Lastly by clicking on the main image, the tool cycles to a second algorithm, which produces an alternative, black and white, visualization in parallel.

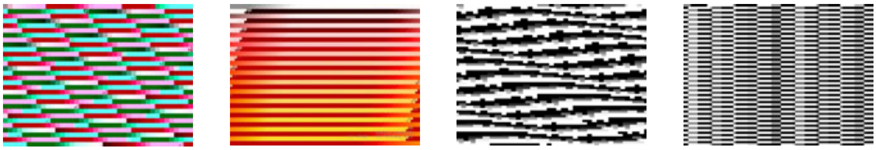


V_1 V_3

Fig. 5. This shows the progression of the *TEXT2IMAGE* interface from its initial release (V_1) to its latest refinement in version 3.0 (V_3)

4.3 Evaluating the New in Precisely Mishandled Images

Aspects of the digital images created through *TEXT2IMAGE* tend to be reminiscent of the 8-bit graphics found in video game consoles from the early 1980's. Where this deviates is in the complex graduating hues and progressive repeating structures that compose these block-like pixels (see Fig. 6). Between the two algorithms available within the tool, the first one returns a greater surprise in the variety of immediate visual output through the wide variations in color values. The output of the second algorithm is seemingly less varied, albeit it contains a more complex visual structure. Optical illusions of three-dimensional depth are often experienced through an adjustment of scale and distance from the image.



alg_01$_1$ *alg_01*$_2$ *alg_02*$_1$ *alg_02*$_2$

Fig. 6. A sample view of output images from both the color algorithm (*alg_01*) and the black and white algorithm (*alg_02*)

Throughout the research phase leading up to *TEXT2IMAGE*, a frequently asked question was how the newly generated images could be graphically applied. As the website began to be covered in blogs and traffic surged, the concern was drowned out by a general attraction in both the technique of creation and the resulting imagery. This notion was well observed as one particular blog wrote, "*text2image is one of those 'data art' works that balances on the line of total uselessness and visual fascination. While there is no immediate purpose, one still spends minutes playing around with it.*" [14]. In contrast it was also described as a medium in its own right for artistic expression, "*Watercolor, oil paints, charcoal, sanguine, India ink. . . There exists a large palette of techniques for those who want to unleash their creativity and create an image. Ted Davis adds a new one: the Web!*" [15].

Hypothetical applications for the imagery were continually pursued, including re-upholstered Ikea furniture catalogs, personalized apparel and revised visualizations for cookbook recipes. Nevertheless it was a half-year later, as a design student from the UK used *TEXT2IMAGE* to generate book covers for a design contest entry [16], that the produced imagery was once again dispersed across a new range of blogs. Migrating now from visualization and technology based blogs to those in the realm of graphic design, a confirmation grew in that the generated imagery contained previously unforeseen qualities.

5 Conclusion

Under the premise that nearly all of our digital files are speaking the same language of binary code, it became plausible to mix such contrasting phenomena as text and image together. In an analog domain, this could be done through the use of collage, physically layering one over the other. However in the digital domain, these elements have the capacity of being combined underneath the surface, on the structural level. By mixing the contents of a *cross-media visualization* was created. Through the *precise mishandling* of each file format, its requirement to act as a transparent medium is disregarded, as the inherent visual qualities it wishes to share are further explored and exposed.

As new visual discoveries were made through the introduction of forced glitches, a lack of proper tools to accommodate this practice was quickly realized. Ten years ago at the time of Kim Cascone's seminal text on the use of failure and glitches in electronic music [17], a variety of tools for the emulation and implementation of audible glitches already existed. A call was made in the concluding paragraph for the development of tools that would help to educate the musician in this field of practice, specifically beyond a "blind experimentation". Following on the heels of *glitch music*, *glitch art* [18], as it has grown to be known, severely lacks the tools for both exploration and education in usage of the glitch. While the hex-editor continues to be a usefully appropriated tool in order to accomplish both of these tasks, it is in dire need of specialized functions relative to the file types that visual media embodies. This inspired the development of *HEADer_REMIX* [19], an alternative hex-editor, limiting access to the safely malleable regions within the image header, as well as providing an instant visual update from the resulting changes to the pictorial surface. What it lacks in novice approachability, it makes up for through education in regard to the cause and effect relationship of the header and image surface. Perhaps one day, professional software will offer a sort of *digital native* user access, allowing for the selective disabling of error correction techniques. As glitches and bugs find their way into our working documents, an awareness for the media will grow, parallel to the development of a new visual language—as discovered through *TEXT2IMAGE*—accompanying our journey further into the realm of digitalization.

References

1. The Playboy Interview: Marshall McLuhan (1969), `http://www.nextnature.net/?p=1025`
2. Manovich, L.: The Language of New Media, p. 289. MIT Press, Cambridge (2001)
3. Cascone, K.: The Aesthetics of Failure: "Post-Digital" Tendencies in Contemporary Computer Music. Computer Music Journal 24(4), 12–18 (2000)
4. Goriunova, O., Shulgin, A.: Glitch. Neural 28, 110–119 (2007)
5. Rhizome blog discussion, `http://rhizome.org/editorial/2009/feb/25/pixel-bleed/`
6. DATA DIARIEZZZZZZ, `http://www.turbulence.org/Works/arcangel/`
7. Columbia University Lecture, `http://www.columbia.edu/itc/soa/dmc/cory_arcangel/`
8. Manovich, L.: The Language of New Media, pp. 45–46. MIT Press, Cambridge (2001)
9. Lister, M., et al.: New Media: A Critical Introduction, 2nd edn. Taylor & Francis, Abington (2008)
10. Wallace, G.K.: The JPEG still picture compression standard. IEEE Transactions on Consumer Electronics 38(1), 18–34 (1991)
11. JPEG Compression, `http://www.fileformat.info/mirror/egff/ch09_06.htm`
12. JPEG, Wikipedia: The Free Encyclopedia, `http://en.wikipedia.org/wiki/JPEG`
13. TEXT2IMAGE, `http://www.text2image.org`

14. Moere, A.V.: Text2image: Transforming Text into Visual Glitches. Information aesthetics (2009), http://infosthetics.com/archives/2009/02/text2image_transforming_text_into_a_glitch.html
15. Meyer, J.: L'art du texte. Le Matin Dimanche, Lausanne (2009), http://www.lematin.ch/multimedia/coup-coeur/art-texte-87810
16. Hello Bauldoff (2009), http://hello.bauldoff.com/post/182036933/experimental-book-covers-for-a-new-line-of/
17. Cascone, K.: The Aesthetics of Failure: " Post-Digital" Tendencies in Contemporary Computer Music. Computer Music Journal 24(4), 12–18 (2000)
18. Moradi, I., Scott, A., Gilmore, J., Murphy, C.: Glitch: Designing Imperfection. Mark Batty Publischer, New York (2009)
19. HEADer_REMIX, http://www.ffd8.org/header_remix/

The Nature of Adaptive Interactions and Potential for Strategic Application

Justin Davis

Madera Labs, USA
justin@maderalabs.com

Abstract. The ability for an interaction with a user to adapt over time can aid in the creation of more engaging user experiences as well as helping to drive business growth through strategic application. These adaptations may take two broad forms: active and passive, with each exhibiting certain characteristics depending on context of use. Each of these models is rooted in the system model for adaptation, a 4-step cycle of information gathering, induction and response that forms the foundation for designing the nature of an adaptive system. In this paper, we describe the nature of adaptation, explore various contexts of use with modern examples and describe potential strategic business application for this interaction methodology.

Keywords: adaptive interactions, adaptation, interaction design.

1 Introduction

Historically, many designed human computer interactions have involved engagement that is primarily static, starting and stopping with the span of time that the user is actively using the software, and failing to transcend that period in a meaningful and engaging manner. Today, with the increasingly ubiquitous nature of computing, nearly infinite data storage scenarios and an array of new interaction modes via hardware and software combinations, this static nature of software begins to reveal its limitations. As new understanding emerges about how humans interact with technology, and how technology can play an ever-increasingly persistent role in their lives, the role of adaptive interactions becomes not only more feasible, but a key player in the design of more engaging, persuasive and delightful human-computer interactions.

The notion of adaptive interactions - particularly with regards to an interface that changes over time according to user behavior - is not novel. Per Langley:

> *An adaptive user interface uses traces of its interactions with users to automatically build a model that influences the systems behavior to better suit the user's preferences. (Langley, 1999)*

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 221–227, 2011.
© Springer-Verlag Berlin Heidelberg 2011

This alteration of the user interface according to patterns of behavior found in data has been explored in scientific projects like INCA (Gervasio, Iba & Langley, 1999) and SYSKILL & WEBERT (Pazzani, Muramatsu & Billsus, 1996), as well as popular internet websites like Amazon and Netflix. In all of these instances, the ability for an interface to adapt to a user causes the computer to deliver a more relevant and targeted response to the user, thereby increasing efficiency and driving loyalty.

For the purposes of brevity, we will use the word "device" to refer to the item the user is interacting with. This is used with the understanding that this may refer to a computer-enabled device, but also to a human-to-human interaction in the case of adaptation in the service design sector.

2 Dimensions of Adaptation

Adaptive interactions require that new dimensions be considered in the design process. Whereas static interactions only require the consideration of the standard three dimensions, designing for adaptation requires the active exploration of at least one additional dimension: time. In addition to time, the dimension of touchpoints might also be considered in more ubiquitous and pervasive applications.

Because adaptation happens by introspecting past performance and altering the state of the object for future performance, the consideration of time is a requirement in the design of adaptive interactions. To design experiences that successful leverage the potential of adaptive interactions, designers must think of time as a primary design constraint and use it to frame the problem space from the outset.

Time is a required component of adaptive interactions, for if there was no record of data over time, there would be no basis on which to craft an adaptation. In addition to time, the dimension of touchpoints (broadly: location) plays an important part in helping the designer to craft engaging adaptive interactions.

From a systemic view, touchpoints represent discreet devices or locations where a user interacts with a device. Salient to this discussion is the requirement that these devices be able to track the user and their behavior, store that data, and use that data across the system for changing the nature of the interaction over time. While not required for the construction of adaptive interactions, the dimensionality of touchpoints does provide the designer with a compelling set of opportunities for extending the interaction past a static object.

3 A System Model for Adaptation

Adaptation can be described as a loop with four discreet action points. This loop constitutes the cycle by which data is collected, analyzed, and acted upon for the purpose of changing the nature of the interaction with the user.

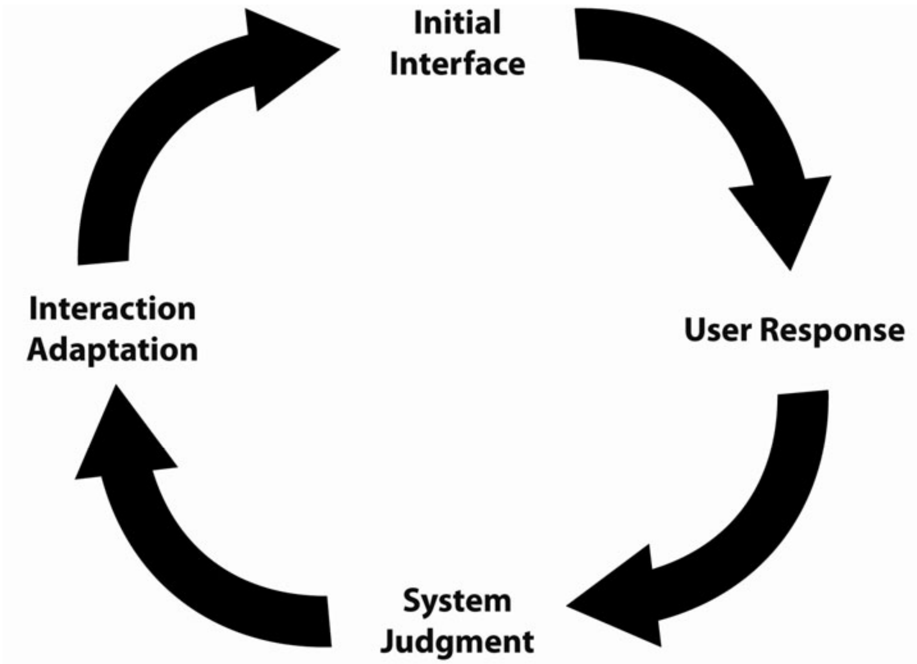


Fig. 1. A System Model for Adaptation

3.1 Intial Interface

The initial interface represents the state of the device before the user interacts with it. This happens both at the start of the engagement with the device - when the user is using the device for the first time - as well as at the completion of this loop, after the interaction has adapted and before the user interacts with the device again.

The initial interface can be thought of as being in a "listening" state, passively waiting for the user to act upon it.

3.2 User Response

Immediately following the display of the initial interface to the user is the user's response, where an action is applied to the interface. For an interaction to be adaptive, this response must be stored for later retrieval and analysis. Per our earlier discussion regarding time, user responses should be attached to a specific timestamp, so that induction may be carried out with the history of responses intact and chronological.

User response may fall into one of two broad categories: active and passive. In an active response (later discussed as an "Active Adaptation"), the user consciously and intentionally acts upon the interface to drive the adaptation, aware of the adaptive nature of the interaction and their relative effect on it. In a passive response (later discussed as a "Passive Adaptation"), the user does not drive the adaptation proactively. In this model, the system collects behavioral data on behalf of the user by evaluating behaviors over time, as opposed to the use of concrete decision points.

3.3 System Judgment

After a user has taken action on the device, the device uses the data collected from this interaction to derive a conclusion about what the user's behavior means. This is an inductive process, and may be handled by a myriad of predictive modeling and machine learning algorithms. Because the focus of this paper is on the design of adaptive interactions - not the specific mathematical formulation of inductive techniques - we will not cover in detail the specific mathematical formulas needed to carry out this induction. Moreover, Langley and Simon (1995) note that the particular induction method is less important than the steps of problem formulation, representation engineering and data collection/preparation.

Salient to this discussion is that the system must use user behavior and/or response data to formulate the basis for adapting the interaction before the next session with the user.

3.4 Interaction Adaptation

During the final phase of this system model, the device must actually execute the necessary alterations to the interaction before the user interacts with the device again, thus starting the loop over. These alterations may be tangible to the user, such as a change to the user interface or information being served to the user, or may be intangible, such as the adjustment of the user model internally within the system. It should be noted that the point of adaptive interactions is to eventually change the tangible interaction with the user, so although a device may spend some cycles of this loop adjusting the user model instead of the interface, the tangible adjustment of the interaction is eventual.

3.5 Two Models of Adaptation: Active and Passive

Two main implementation methodologies exist for adaptive interactions: active and passive. Each of these describes the nature of the user's role in driving the adaptation, and has consequence on the algorithms selected to drive the change, as well as the design of the interface itself.

3.6 Active Adaptation

In an actively adaptive model, the nature of the interaction changes as a result of a conscious and intentional effort on the part of the user to drive this adaptation. These systems actively solicit input from the user, using the user's response to help drive the inductive process for the adaptation.

Actively adaptive interactions display certain characteristics that help to define the model. With an actively adaptive model, the user interface must offer the user decision points to drive the adaptation. These decision points generally take a binary, yes/no form, allowing the user to play a persistent role in crafting the adaptation. These decision points may offer explicit choices that map directly to the construction of the algorithm, or may offer choices that are more implicit in nature. These implicit choices utilize associated metadata to help drive the inductive process. However, regardless of the explicit or implicit nature of the user's input, actively adaptive systems rely on a user driving the nature of the adaptation directly through decision points.

One advantage of actively adaptive systems is the simplicity by which the predictive algorithms may be implemented. Because the system is changing behavior due to a user's overt and explicit instruction, the reliance on predictive modeling algorithms is lessened. From an implementation standpoint, these models "learn" by offering the users choices, observing the choice and making the same choice later for the user. They may also map the metadata from a user's implicit choice to other possible outputs, adjusting the information served to the user according to profiles attached to these decision points. Although more simple to implement, actively adaptive systems require a greater degree of involvement on the part of the user, and may be less compelling than passively adaptive interactions in terms of experience.

Pandora, based on the work of the Music Genome Project, is an example of an actively adaptive interaction. The application works by soliciting input from the user as to what type of music they would like to listen to. This input may be in the form of artist, song or genre. After the user has supplied this data to the application, Pandora proceeds to play music that is similar, based on metadata attached to each piece of media.

While listening, a user is able to adjust the accuracy of the station by giving a "thumbs up" or "thumbs down" to individual songs. This action is the required active input for the active adaptation. After indicating their preference, Pandora adjusts its algorithm by examining the characteristics of that song, and serving up future music based on the user's preferences. This is an implicit decision point, where the user is instructing to the system that they do or do not like a certain song, and using the associated metadata for that song to drive the adjustment of the inductive process. Pandora also implements an explicit decision component in this process, whereby a particular song is no longer played after two “thumbs down” by a user. The difference in implicit and explicit decision points is quite clear here: with an implicit decision point, the user’s feedback is extrapolated to other instances using attached metadata, whereas an explicit decision point instructs the system as to how to behave in the *exact* circumstance in the future (in this case, not playing the exact song that the user has applied the feedback to).

3.7 Passive Adaptation

In a passively adaptive model, the nature of the interaction changes automatically over time, through induction of the user's past behaviors. These systems do not require active solicitation of the user, instead adjusting the interaction automatically in subtle but meaningful ways over time.

Because of the heavy reliance on induction, passively adaptive systems are more difficult to implement, and require more careful examination of the data collection and representation methods used, to assure that the alterations in the model are representative of a user's actual desires.

Passive adaptation also carries with it extra implementation burden not always associated with actively adaptive models. In a passively adaptive model, the potential for confusion and disorientation on the part of the user is greatly increased, due to the fact that the user is not inherently involved in the adaptation of the interaction. As a system adapts to better serve a user, the nature of these adaptations must be smaller and more incremental than in that of an actively adaptive system. Said another way, much more behavioral data is needed to effect changes in a passively adaptive interaction than is needed in an actively adaptive model.

In addition to the potential for confusion and disorientation, passively adaptive interactions must also employ override mechanisms for allowing users to adjust the adaptation manually, or correct an errant assumption made by the algorithm. These overrides act as an adjunctive actively adaptive component, and demonstrate that passively adaptive interactions may not perform well as purely passive, but as a blend between the two modes.

Passively adaptive interactions are becoming more common, especially in the web domain. Google's recent product, Priority Inbox, is an example of such a system, and employs the aforementioned characteristics of a passively adaptive system.

With Priority Inbox, the system attempts to use induction of user behavior to establish priority in an email inbox. It does this by using inductive methods, observing user behavior such as who you email, which messages you open, which keywords appear in messages commonly opened and which messages you reply to (Google). However, as is characteristic of a passively adaptive interaction, the system also features override mechanisms, in the form of buttons that allow the user to manually label a message important or not important. Per our earlier discussion, these override mechanisms are strongly encouraged in a passively adaptive interface to allow the user to adjust the nature of the adaptation, should induction prove inaccurate.

3.8 Adaptive Interactions as a Strategic Business Tool

The use of adaptive interactions can be leveraged as a significant driver of business in a transactional situation. Adaptive interactions have the ability to deliver value to a user in more poignant and relevant ways, which can drive revenue through an increase in sales conversion and by intentionally raising the switching costs associated with leaving a provider.

Amazon has demonstrated the efficacy of adaptive interactions as a strategic business tool. By using product recommendations – using adaptive models such as content-based and collaborative filtering (Langley, 1999) – Amazon has increased revenue by increasing the likelihood that that users will convert, due to an increased relevancy of the information being displayed to them. These recommendations have accounted for up to 35% of Amazon's yearly revenue (VentureBeat), demonstrating the power of adaptive interactions for driving revenue-generating conversions. The adaptation models used by Amazon to display recommendations are of the passive class, learning from user behavior over time to display more relevant content in future interactions.

In addition to increasing conversion, adaptive interactions can also drive revenue creation by intentionally raising the switching costs associated with leaving a provider. Burnham, Frels and Mahajan (2003) describe personal relationship loss costs as losses that may be incurred when a bond is broken between a customer and the people with whom that customer interacts. This relationship may be bolstered or defined by a rich data set and existing history shared between customer and provider, which allows the provider to deliver a more relevant experience through the use of adaptive interaction. Should the customer leave a provider with which they have a long history, they lose the benefit and comfort associated with the former provider.

This increased switching cost speaks to the benefit of leveraging user data to create more relevant experiences, exposing the potential switching costs to the user and driving deeper loyalty with the provider.

4 Concluding Thoughts

Clearly, the ability for an interaction to adapt to a user carries with it both compelling experience advantages as well as business growth opportunities. The nature of adaptation is complex, consisting of varying interaction methodologies, inductive algorithms and strategic business considerations. Despite this complexity, the potential of adaptive systems to deliver a more engaging and relevant experience to users is exciting. To be successful, a designer must possess an understanding the user's experience over the dimensions of time and touchpoint, and understand how to design each discreet moment throughout the system model for adaptation. In the application of adaptive interactions, the devil is in the details; care must be taken to carefully consider the correct adaptive scenario to be used, and each component of the adaptive interaction must be designed with caution and intent. Executed successful, however, adaptive interactions can create exciting new moments for people interacting with technology and can aggressively drive business campaigns.

References

1. VentureBe, `http://venturebeat.com/2006/12/10/aggregate-knowledge-raises-5m-from-kleiner-on-a-roll/`
2. Burnham, T., Frels, J., Mahajan, V.: Consumer Switching Costs: A Typology, Antecedents, and Consequences. Journal Academy of Marketing Science 31(2), 109 (2003)
3. Gervasio, M.T., Iba, W., Langley, P.: Learning User Evaluation Functions for Adaptive Scheduling Assistance. In: Proceedings of the Sixteenth International Conference on Machine Learning, Bled, Solvenia, Morgan Kaufmann, San Francisco (1999)
4. Google, `http://mail.google.com/support/bin/answer.py?hl=en&answer=186543`
5. Langley, P.: Machine Learning for Adaptive User Interfaces (1997)
6. Langley, P.: User Modeling in Adaptive Interfaces. In: Proceedings of the Seventh International Conference on User Modeling, Banff, Alberta. Springer, Heidelberg (1999)
7. Pazzani, M., Muramatsu, J., Billsus, D.: SYSKILL & WEBERT: Identifying interesting web sites. In: Proceedings of the Thirteenth National Conference of the American Association for Artificial Intelligence, pp. 54–61. AAAI Press, Portland (1996)

Studies on Imaging Methods to Realize Effective BCI through ERPs

Mariko Funada[1], Yoshihide Igarashi[2], Tadashi Funada[3], and Miki Shibukawa[4]

[1] Hakuoh University, Department of Business Administration, Oyama, Tochigi, Japan
[2] Gunam University, Professor Emeritus, Kiryu, Gunma, Japan
[3] Rikkyo University, College of Science, Toshima-ku, Tokyo, Japan
[4] Hakuoh University, Department of Education, Oyama, Tochigi, Japan
mfunada@fc.hakuoh.ac.jp.jp

Abstract. This paper is to clarify the usefulness of imaging methods, such as imaging figures or words, for realization of effective Brain Computer Interaction (BCI). We design some experiments in which subjects engage in imaging figures or words. Event Related Potentials (ERPs) are measured through the experiments. Some variables of ERPs are introduced for the data analysis of our experimental. We consider ERPs as vectors of various components. Using the norms of the vectors, we discuss which imaging methods are suitable.

Keywords: BCI, Event Related Potential, vector, norm, analysis of variance.

1 Introduction

In order to realize effective BCI (Brain Computer Interaction) we are required to devise suitable imaging methods. ERPs (Event Related Potentials) [1] [2] are obtained by averaging EEGs (Electroencephalograms) that can be considered the reflection of mental pictures occurred in the brains. In particular human beings have a tendency to visualize various matters including their thoughts and knowledge. Our assumptions concerning this tendency are as follows:

- When a subject is taking a multiple-choice test, he/she may visualize a figure (number) or a word that corresponds to his/her choice. We assume that ERPs as the reflection of the visualization for the figure may be somewhat different from ERPs as the reflection of the visualization for the word.
- We assume that for a multiple-choice test, the test results may be influenced by the way of visualizing his/her choices.

In this paper we try to confirm these assumptions by experiments and data analysis.

2 Methods

2.1 Experiments

1) The subjects: Eleven students took part in the experiments. They are nine right-handed men and two women, 20 to 22 years old. We use lower-case Roman letters from *a* to *k* to identify them.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 228–236, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2) The place of the experiments: The laboratory of the first author at Hakuoh University.
3) Stimuli: We use 20 kinds of stimuli (see Fig.1). A stimulus shown in the display is a Chinese character together with three choices for its pronunciation. Each stimulus is displayed in the square of 3.28 cm×9.68 cm. We call each of these stimuli a question.

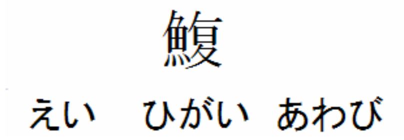


Fig. 1. An example of a stimulus (question)

4) Tasks: We use three task types called *A*, *B*, and *C*. In experiments, a subject watches a Chinese character, and then he/she chooses its proper pronunciation from the three choices. The leftmost choice, the middle choice and the rightmost choice correspond to numbers 1, 2, and 3, respectively. The subject images figure (number) 1, 2 or 3 that corresponds to its proper pronunciation. For example, if the subject chooses the leftmost one, then he/she images figure 1 and read it silently in Japanese (1 is pronounced “ichi” in Japanese). We call this task type *A*. When the subject images one of the three words, *leftmost*, *middle* and *rightmost* as the proper pronunciation and reads it silently in Japanese, the task is called type *B*. When the subject simply inputs a number from 1, 2 and 3 without considering its image as the proper pronunciation, the task is called *type C*.
5) Display of stimuli: a sequence of stimuli is displayed in a CRT (Cathode Ray Tube) of 19 inches placed in front of the subject. A sequence of 40 stimuli is called a set of stimuli. Two sets of stimuli are executed in one day. For any stimulus displayed in the CRT, the subject watches it without moving his/her eyes.
6) Time duration for the display of a stimulus: Each stimulus, as shown in Fig. 1, is displayed for 1 second. The interval between two consecutive stimuli is randomly chosen in the range from 400 [ms] to 600 [ms].
7) Time duration of an experiment: About 1 minute is spent for a set of stimuli. A subject takes a minute interval between two consecutive sets of stimuli. Consequently, the time duration of 2 sets of stimuli for all tasks excluding the interval time is about 6 minutes.
8) EEGs: Single polar eight channels of “International 10-20 methods” are used for measuring EEGs. The measurement positions are at Fp_1, Fp_2, C_3, C_4, O_3, O_4, C_z, and P_z. The base is at A_1 that is connected to A_2. In this paper, we analyzed the data measured at C_3, C_4, C_z, and P_z.
9) The sampling frequency for A/D: 1 kHz.

2.2 Analyses

We process the recorded EEGs to obtain ERPs in the following way:

1) The recorded EEGs are filtered by an adaptive filter.
2) The filtered data are normalized by the average and the standard deviation of the data.

3) The selecting and averaging method is to divide the normalized EEGs by a threshold $L(=0.5)$ into three classes I, II, and III [3], [4]. Then we obtain the three classes of ERPs. We call this method the "selecting method".
4) Several variables are defined in order to characterize the ERPs.
5) Two-way-layout analyses of variance are applied to the factors of ERPs; positions of electrodes, and imaging matters using the variables defined in the paper.

We mainly use ERPs in class I to examine the two assumptions described in Introduction.

3 Results

3.1 Recorded Data, and Filtered and Normalized Data

Examples of EEGs measured from subject *b* are shown in Fig. 2. The time elapse [ms] from a given stimulus is shown on the horizontal axis. The amplitude of measured data is plotted in the vertical direction. In Fig.2, the lowest, the second lowest, the third lowest, and the highest waveforms are plotted from the data measured at C_z, P_z, C_3 and C_4, respectively. The measured data contain 50 [Hz] noise and other types of noises as well. These noises would be caused by electromyography, blinks or body movements of the subject and others. Before starting the experiments, each subject is asked that he/she should make an effort to minimize his/her blinks and body movements. Noises of higher frequency than frequency of EEGs are also minimized. The recorded data are filtered and normalized. These data are given in Fig. 3.

3.2 ERPs Obtained by the Averaging Method and the Selecting Method

ERPs are calculated by the averaging method. ERPs in Fig.4 are an example of them. ERPs are also calculated by the selecting method. ERPs in Fig.5 are an example of them, and they are calculated from the same data given in Fig.4. Amplitudes of ERPs in Fig.5 are larger than those in Fig.4. The peaks of amplitude appeared in Fig.5 are clearer than those in Fig.4. Almost the same results are obtained from ERPs of other subjects. We use the ERPs obtained by the selecting method for the purpose of our analyses.

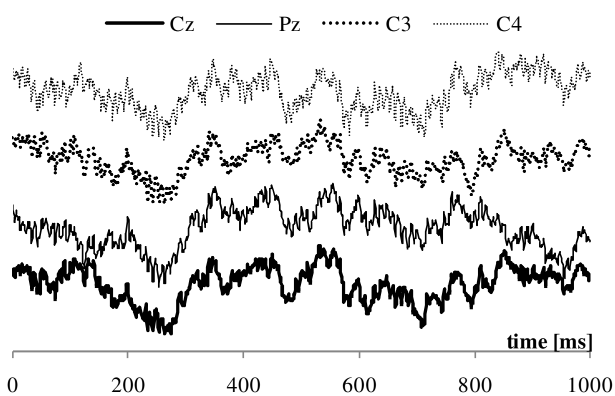


Fig. 2. An example of recorded data (Subject: *b*, Task: type *A*)

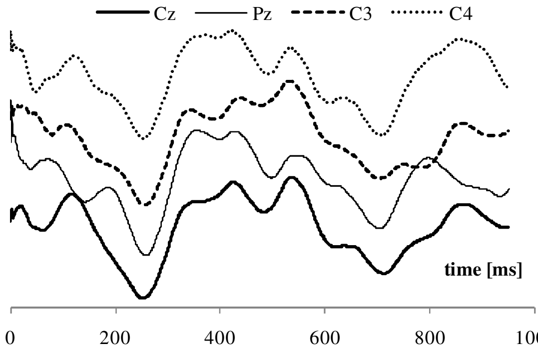


Fig. 3. The data given in Fig. 2 are filtered and normalized (Subject: *b*, Task: type *A*)

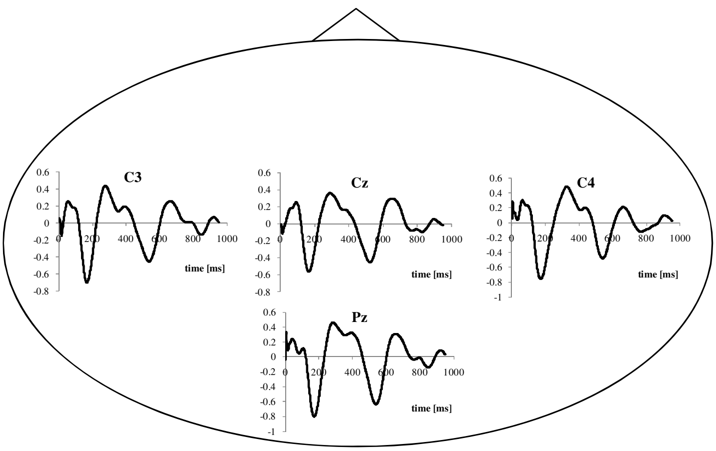


Fig. 4. An example of ERPs obtained by the averaging method (Subject: *d*, Task: type *A*)

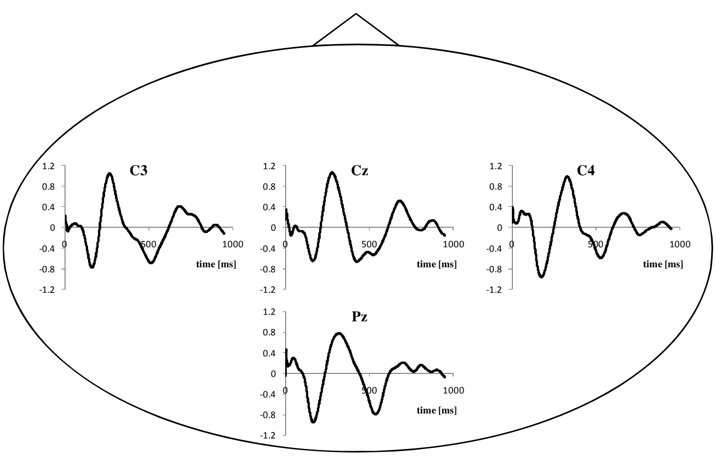


Fig. 5. An example of ERP obtained by the selecting method (Subject: *d*, Task: type *A*)

3.3 Comparison among Task Types

As shown in Fig.6 (Subject: *b*) and Fig.7 (Subject: *a*), we notice that there are somewhat differences among the waveforms of ERPs for task types. The differences among the ERPs in Fig.7 are more significant than those of the ERPs in Fig.6. In Fig. 6, the positive peaks are indicated by P1 and P3, and the negative peaks are indicated by N2 and N4. In Fig.7, we notice that the latency of the peak P3 for task type *B* is shorter than those of the peaks P3 for other task types. Clear differences can be recognized among the waveforms of ERPs for the task types.

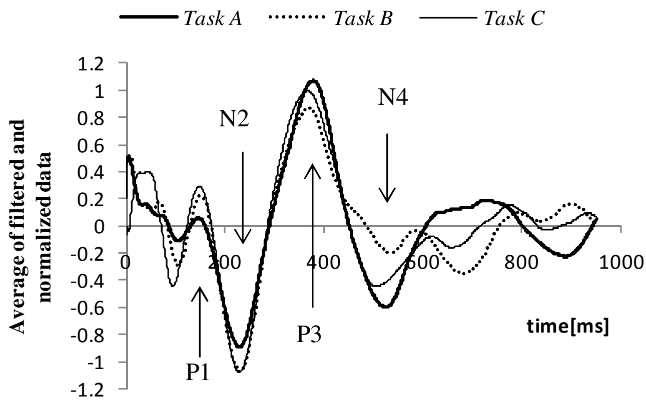


Fig. 6. An example of ERPs showing less significant differences among the task types (Subject: *b*)

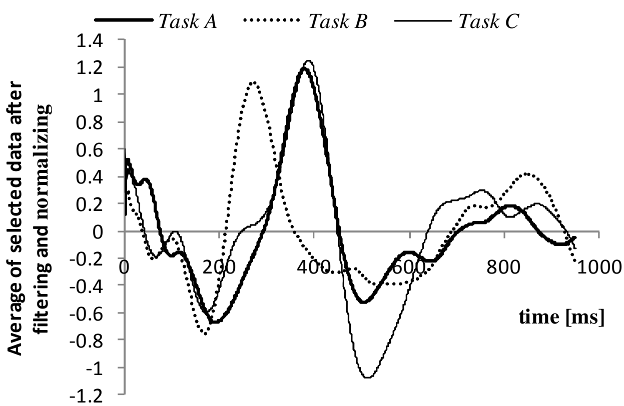


Fig. 7. An example of ERPs showing significant differences among the task types (Subject: *a*)

3.4 Comparison of ERPs between Task Types A and B

In Fig.8 and 9, we show examples of ERPs obtained by task types *A* and *B*. In Fig. 8, we can recognize somewhat differences among the three waveforms when subject *b* is imaging figures *one*, *two*, or *three* for task type *A*. Concerning the amplitude values at peaks P1, N2, P3 and N4, their differences among the three figures imaged by subject *b* are less significant. Three waveforms shown in Fig. 9 are ERPs when subject *b* is imaging "*left*", "*middle*" and "*right*" for task type *B*. We can recognize somewhat differences among these waveforms at all the peaks. The differences among imaging words "*left*", "*middle*" and "*right*" for task type *B* are somewhat more significant than imaging figures "*one*", "*two*" and "*three*" for task type *A*.

In comparison with task type *C*, in Fig. 10 we show the waveforms of ERPs for task type *C*. The amplitudes at P1 and P3 have relatively large differences among "*input* 1", "*input* 2" and "*input* 3". The differences of latencies at N4 among these inputs can be also recognized.

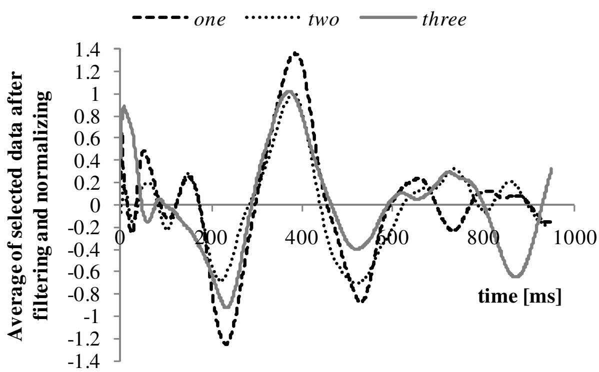


Fig. 8. Examples of ERPs obtained by imaging *one*, *two* or *three* during task type *A* (Subject: *b*)

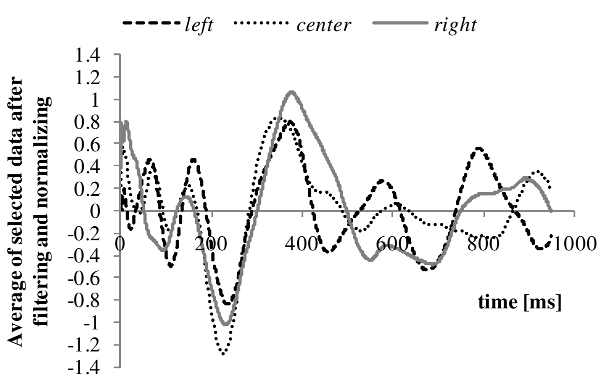


Fig. 9. Examples of ERPs obtained by imaging *left*, *middle* or *right* in task type *B* (Subject: *b*)

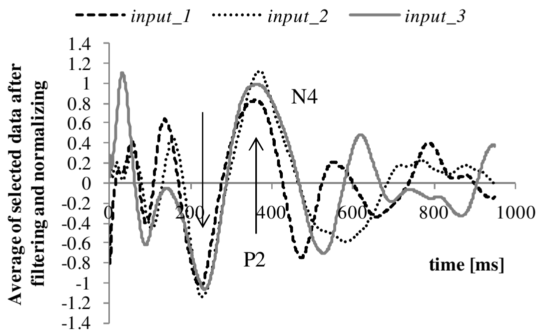


Fig. 10. Examples of ERPs obtained by "*input* 1", "*input* 2" and "*input* 3" without considering any images of figures and words for task type *C* (Subject: *b*)

3.5 Variables and Analytical Results

In Fig. 11, we define ten variables to explain the differences among ERPs in Fig.8, 9 and 10. Variables *P1*, *N2*, *P3*, and *N4* are the latencies of the peaks, and variables *AP1*, *AN2*, *AP3,* and *AN4* are the amplitudes of the peaks. Variables "*area1*" and "*area2*" are the areas of the triangles P1_N2_P3 and N2_P3_N4, respectively.

For each subject, we calculate the values of all the ten variables from his/her ERPs. "Two-way-layout analyses of variance" is done for all the variables. Two factors for "two-way-layout analysis of variance" are used. These are the position of the electrodes and the imaging or inputting matters. The level of the former factor is 4; C3, C4, Cz, and Pz. The level of the latter factor is 3; imaging "*one*", "*two*" or "*three*" for task type *A*, imaging "*left*", "*middle*", and "*right*" for task type *B*, and inputting "1", "2", and "3" for task type *C*. As shown in the left three columns of Table 1, at about 0.05 amplitude level 0.05, there are significant differences among 3 imaging levels only in a few variables. As shown in the right three columns of Table 1, there are also significant differences among 4 levels of electrodes at about 0.05 amplitude level, only in a few variables. These results do not answer to the question which task is suitable to detect imaging matters.

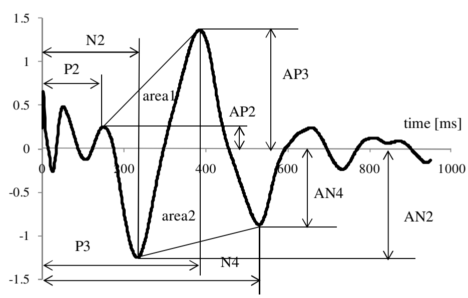


Fig. 11. The definition of variables

Table 1. Analytical results of the variance of a distribution

	factor (imaging or inputting matters)			factor (kinds of electrodes)		
variables	Task *A*	Task *B*	Task *C*	Task*A*	Task *B*	Task *C*
N4				*		
AN2			*			
AP3						*
AN4		*	*			
area2			*			

4 Discussions

As we described in the previous section, the variables do not work to extract clearly the differences among ERPs caused by imaging figures or words. In this section we introduce vectors as shown in formula (1), and their norms as shown in formula (2). In the formulae, index "*i*" indicates task type *A* or *B*, and index "*j*" indicates one of three cases; imaging "*one* ", "*two*", and "*three*" for task type *A*, and imaging "*left*", "*middle*" and "*right*" for task type *B*. $ERP_{ij}(t)$ is the amplitude at t [ms] for task "*i*" and imaging "*j*".

$$\mathbf{d}_{ij} = (ERP_{ij}(t), ERP_{ij}(t+1), \cdots, ERP_{ij}(t+m)) \tag{1}$$

where $i = A, B,\ j = 1,2,3,\ \ t = 100,\ \ m = 600$.

$$D_j = \|\mathbf{d}_1 - \mathbf{d}_2\| + \|\mathbf{d}_2 - \mathbf{d}_3\| + \|\mathbf{d}_3 - \mathbf{d}_1\| \tag{2}$$

We calculate all the norms, and show in Table 2 which task type is more significant.

Table 2. Task types with larger norms D_j than the other task type

Subjects	C3	C4	Cz	Pz	Subjects	C3	C4	Cz	Pz
a	***B***	***B***	***B***	***B***	*h*	***B***	***B***	***B***	*A*
b	***B***	***B***	***B***	***B***	*f*	***B***	*A*	***B***	*A*
i	***B***	***B***	***B***	***B***	*j*	*A*	***B***	*A*	***B***
k	***B***	***B***	***B***	***B***	*e*	*A*	*A*	*A*	***B***
c	***B***	*A*	***B***	***B***	*g*	*A*	*A*	*A*	*A*
d	***B***	*A*	***B***	***B***					

The results explain which task reflects more in the differences among subjects. In the most case (7/11 63.6%), imaging word is more suitable than imaging figures to detect which one is chosen from the three choices. In several cases (2/11=18.2%), imaging figure is more suitable for the same purpose. In the rest of cases (2/11=18.2%), we cannot recognize any clear differences between words and figures.

5 Conclusions

From our experiments and statistical analysis, we can say that the selecting method is valid to extract useful ERPs. In our data analysis, we categorize ERPs into three classes; data with positive and large amplitude, data with low amplitude, and data with negative amplitude. Two assumptions were examined. One is that visualization of figures (numbers) or words in the brains would be differently reflected in ERPs. The second one is that for a multiple-choice test, the test results may be clearly influenced by visualizing way of his/her choice. From our statistical analysis, the first assumption was positively confirmed. Concerning the second assumption, we could not obtain any clear evidence that it is positively confirmed since the differences among the effects of imaging matters are much dependent on individual subjects. It would be an interesting problem how we could overcome this difficulty.

References

1. Picton, T.W., Bentin, S., Berg, P., Conchin, E., Hillyard, S.A., Johnson Jr., R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M., Taylor, M.J.: Guideline for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37, 128–152 (2000)
2. Luck, S.J.: An Introduction Of The Event-Related Potential Technique. The MIT Press, Cambridge (2005)
3. Funada, M., Shibukawa, M., Igarashi, Y., Funada, T., Ninomija, S.P.: Some characterizations of event-related potentials by reconstructing the distribution of the potentials. Journal of Human Factors 46(2), 144–156 (2010)
4. Funada, M., Igarashi, Y., Funada, T., Shibukawa, M.: A Method For Extracting Meaningful Signals From Event Related Potentials. In: The Eighth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications, Innsbruck (2010) (in printing)

Multitouch Tables for Collaborative Object-Based Learning

Jacob George[1], Eric de Araujo[2], Desiree Dorsey[3],
D. Scott McCrickard[1], and Greg Wilson[1]

[1] Department of Computer Science and Center for Human-Computer Interaction, Virginia Tech, Blacksburg, VA, USA
[2] Department of Information Technology and Department of Philosophy, Bethel College, Mishawaka, IN USA
[3] Department of Psychology, Bennett College, Greensboro, NC USA
{jacobgc,mrwilson,mccricks}@vt.edu,
eric.dearaujo@bethelcollege.edu,
ddorsey@bennett.edu

Abstract. Multitouch technology on tabletop displays allows children to interact with digital objects in collaborative activities. This paper explores both evolutions in hardware and opportunities in software toward supporting the engagement of children, with consideration of impact on user interfaces. We outline a demonstration of our Multitouch Education Table (MET), a portable hardware system and virtual card game suite targeted for use by elementary school students.

Keywords: multitouch, tabletop, children, education, evaluation.

1 Introduction

Multitouch systems make possible multiple simultaneous inputs not only from a single user but from multiple users. Chronicled in a report from Buxton [1] and popularized in Jeff Han's 2005 UIST paper [2], multitouch is now widely found in devices from Apple, Microsoft, Google, and others. Researchers and practitioners have developed image manipulation, business applications, games, and, most recently, educational applications (e.g. [3]) to take advantage of the technology. This paper explores the promise that lower prices and greater portability can lead to development of tools for early education.

We discuss hardware features that can yield a low-cost portable tabletop—particularly important in educational settings when cost is a major issue. We describe the advantages to a do-it-yourself approach that leads to modularity and lightweight hardware solutions, and we demonstrate how that can facilitate the development of multitouch applications for student use.

To demonstrate the potential of multitouch in the classroom, we show activities that make use of multitouch digital cards. The card paradigm supports real-world card behaviors (e.g., moving and flipping) but also virtual-world behaviors like linking and mapping.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 237–246, 2011.
© Springer-Verlag Berlin Heidelberg 2011

We provide details about a lab experiment that compared multitouch and traditional desktop platforms. Performance metrics were collected for linking tasks, including time to make a match, number of correct responses, and number of errors. In addition, participants took pre-test and post-test surveys to measure knowledge, preference, and additional feedback.

Finally, our paper also reports on reactions of children to multitouch technology. The possibility for multiple students to interact with the multitouch system might prove more cost effective than providing each student access to an individual computer and support collaborative educational goals. As this type of technology makes its way into the classroom, it holds potential to facilitate “learning by doing” for groups of children [3] through support for object-based manipulation.

2 Customizing Hardware for Education

Multitouch table configurations vary in shape, size, and construction. Two commercial examples include the Microsoft Surface and Smart Technologies' SMART Table. The former is used in a variety of environments, while the latter is specifically designed for education. However, these systems face barriers to adoption in educational settings because of their low portability, lack of modularity, and high cost [4]. The goal of MET's hardware design was to provide a portable, modular, and low-cost multitouch display.

The portability of MET is important in an educational setting where the table may need to be shared between classrooms and buildings. To achieve a highly mobile system the MET is built of lightweight materials such as hard-board panels, wooden legs, and a wooden base. These pieces along with the acrylic top, and electronic components (computer, projector, and camera) collectively weigh 70 pounds and can be transported individually. This makes it more portable than the 150-180 pounds seen in most commercial products (which must be transported as a single unit).

The ability to disassemble the unit contributes to its modular design where individual components can be replaced; for example, a single table frame in a school system could utilize projectors already present in the classrooms of each school. This would eliminate the need to transport a projector when bringing MET to the schools, allowing it to be more easily shared among classrooms. (See Fig 1.)

The Do-It-Yourself (DIY) approach that helped make MET portable and modular also produced it for under $3,000. This makes it a cost effective alternative to the $6000 to $12,000 (or more) commercial alternatives. In addition to the reduced cost and improved portability and modularity, improvements in projector and camera technology—with smaller throw distance—could yield a table smaller than the current 45 inches and lighter than the current 70 pounds. LCD displays instead of a projector hold further benefits.

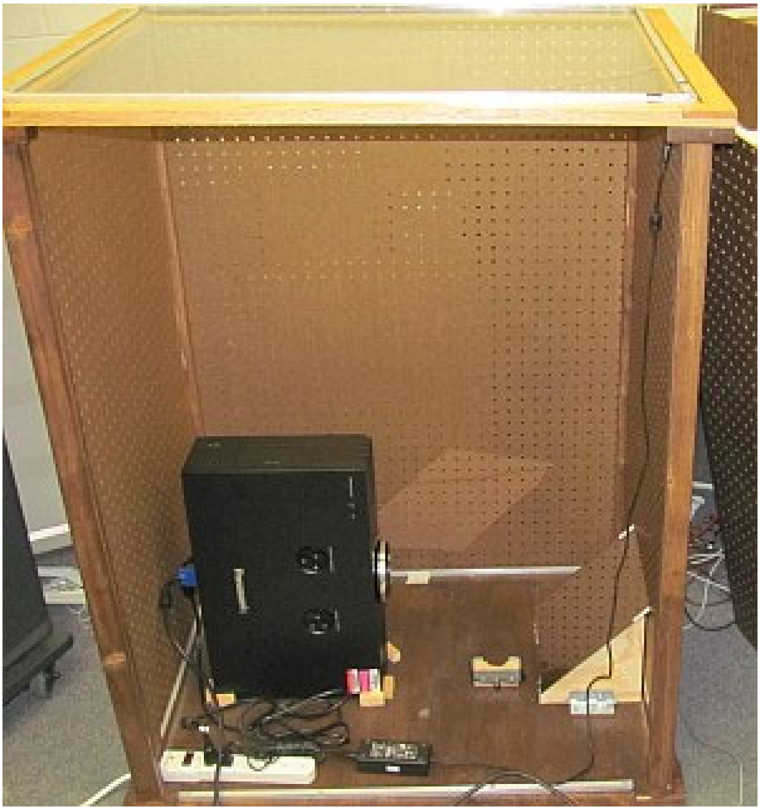


Fig. 1. A side view of the open MET displaying its internal components, including the projector, wiring and power sources, and mirror

3 Digital Card Activities

The decrease in price and increase in availability of multitouch tables and similar technologies has resulted in an increase in software for educational settings, including efforts by and for Microsoft Surface, SMART Table's Interactive-Learning Center, Infusion's Education Suite, and other commercial and academic efforts. Our attempts have focused on *digital cards*, virtual knowledge artifacts that can be directly manipulated by users. Prior work suggests that multitouch artifact-based interactions will result in more interaction and fewer negative interactions [3]. We seek to reinvent experiences with physical cards and single-touch applications, but in so doing, we must be careful not to duplicate the paradigms of physical cards or single touch—respecting the novelty of the platform. Like physical cards, they can be moved, flipped, and can include a combination of text or images. Our MET games include pairing and grouping activities targeted to support children's desires to both collaborate and compete, and that support teachers' need to tailor the applications for their topics.

Our extension of the card paradigm provides unique advantages over physical cards. By storing the content of the cards in an XML document, the learning activities have been abstracted from the specific educational content. This flexibility supports the use of these cards for teacher-selected content that could be modified by entering *topic elements* (e.g., animal-country pairs, or animal life-cycle sequences like egg-larva-pupa) and *relationship types* (explained next) describing associations between elements.

Four relationship types, represented by similar multi-handed gestures that seem to come naturally to young children, can support a wide variety of learning applications. *Linking* connects cards that share an unordered relationship (e.g., animals and countries, or inventors and inventions), accomplished by dragging cards so that they overlap. *Combining* builds upon linking in that cards with a certain relationship are brought together to form a new card (e.g., combining colors like red and blue to create purple). *Mapping* associates objects with a position or region on the screen, like objects to a map or other fixed category. In *sequencing,* the left-to-right or top-to-bottom ordering of objects has meaning, important for subject areas like animal life-cycles and ordered historical events. See Figure 2 for an example of the "link" behavior.

We have built numerous card-based applications for the MET, several of which are highlighted in the remainder of this section. The *Math Match* game, targeted for students in grades 7-12, provides an environment in which students learn about concepts like order of operations and algebraic concepts. Participants in the game can play by themselves, compete with others, or cooperate toward finding the right answer to problems. The *Animal-Continent* game, targeted at young elementary-school-age student, asks participants to match animals to the continent in which the animal is naturally found (see Figure 2). When a card is released on the correct continent on a map, it shrinks and locks into place. The *Inventors-Inventions* game is similar to the Animal-Continent game in that matched pairs are created. Different card sets can target different disciplines (e.g., math, sciences, engineering) and different demographics (e.g., African-Americans, women). Other games allow participants to create matchings, links, or orderings among colors and shapes.

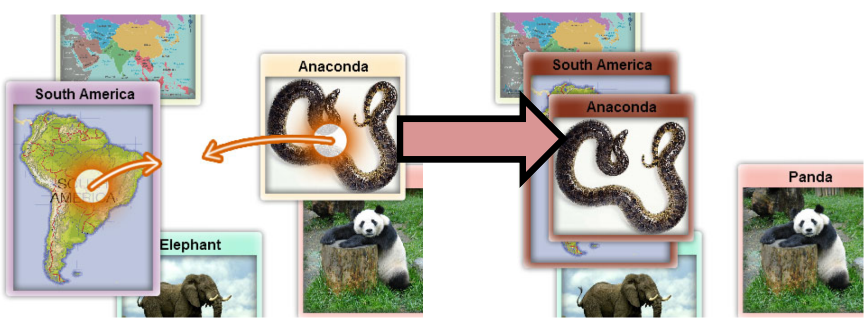


Fig. 2. South America and the Anaconda card are dragged toward each other (left), then when the cards overlap, they "link" and provide visual and auditory feedback (right)

4 User Study: Multitouch Tables and Desktop Displays

A user study was performed to explore performance and enjoyment of the multitouch table environment in a controlled lab-based setting. This study compared performance and preferences for 11 college undergraduate students on three card game activities using our multi-touch table (MET) and using a traditional desktop computer. This user study appears in full detail in another paper [5]; the approach and key results are summarized here.

The students were separated into two groups, and they were asked to work individually on a series of three activities: a color-matching activity, a shape-matching activity, and an inventor/invention-matching activity. Students in one group started with the desktop activities first, and the other group started with the MET. Students in both groups alternated between the desktop and MET until they completed two trials for each activity type on each system. Statistics like the total game time and amount of time the student to complete each individual match were stored.

Among the results, one stood out: users took significantly less time to complete activities using a traditional computer desktop than the MET (p=0.0145). However, user comments comparing the MET and desktop games showed that people preferred using the MET, finding the game more enjoyable on that platform and expressing a desire to see it in classroom settings. It is possible that the lack of familiarity and other technical difficulties and limitations degraded performance with the MET, and that with practice the participants could perform as well or better with the MET as with a traditional desktop system. In addition, technological advances in multitouch surfaces may lead to increased performance.

5 Multitouch in the Field: Notes from Visits with Student Groups

To understand the needs and desires of our target user population, we frequently conduct presentations and hands-on demos with students and educators. It was important to test the durability and portability of the table (to see if it could be disassembled, moved to a different location, and reassembled with ease), test the usability of the card matching applications, as well as interview and gauge the interest of the children in using a multitouch table and possible applications. Figure 3 is illustrative of our experiences in the field.

The first experience with demoing the table came at the ***Destination ImagiNation Regionals*** in Salem, Virginia. Destination ImagiNation is a non-profit organization that features teams of students ranging from elementary school to college competing in creative problem solving challenges. This description made this group a perfect setting to setup the table. The table was set up as a side attraction for the elementary children to experience in between their challenges. No applications had been created at the time, so small demo applications were run instead. These include a chalk board application and a basic visualization application where the children could see abstract visualizations wherever they touched on the surface.

Fig. 3. Typical behavior among children using the MET, with multiple people, hands, and fingers all manipulating digital objects simultaneously

Because this was the first time demoing the table in public, there were some minor obstacles to overcome. There was an issue of finding a good position for the table, because it had to be located near a power source. The table also could not be directly under a light source because it interfered with what the table processed as touches. Our two options became setting up in a high traffic where it was hard for anyone to use the table or further down the hallway where it would not be immediately seen by guests. Although we were not in an ideal location, there were still 10-12 kids that visited the table. They enjoyed playing the demos on the table and stayed at the table for approximately 15-20 minutes. Because there was only one researcher there who was tasked with making sure the table function properly, there was not a chance to interview the children more formally.

The summer camp visit was conducted at the ***STEM Fun! Camp*** hosted by the Institute for Advanced Learning and Research in Danville, Virginia. During the camp, elementary school aged children are exposed to many new technologies and participate in science experiments with the goal of getting them interested in and motivated about STEM (Science, Technology, Engineering, and Math) subjects.

The visit consisted of a formative intervention with 14 participants in the camp, who explored a classification (linking) activity. The participants were brought to the multitouch table in pairs. To begin the session, the participants were given an explanation on what a multitouch table was. They were given a high level overview of how the multitouch table technology worked while being shown the inside parts of the table. The participants were then able to touch the screen and view a visual representation of how the table was sensing and responding to their touches. The children then worked together to match sets of seven or eight animals to their countries or continents of origin (e.g., kangaroos to Australia)—which they were able to complete in about a minute per set, with minimal assistance (shown in Figure 2). Afterwards, the children participated in a short discussion in which they were asked if they thought the multitouch table would be useful in their classrooms and if so, what school subjects would it be most beneficial.

The students reported only minor problems using the technology (primarily related to the need with our technology to touch with fingers or hands and not fingernails), and almost all would be willing to use the table on a frequent and regular basis as part of their learning activities. One issue with the table is that it was frequently too tall for the participants to be able to see and touch the whole surface. To rectify this, we had to use chairs for the children to stand on so that they could have access to the entire table. Obviously, this is not the safest approach to this issue. For future prototypes, we would have to take this issue into consideration and lower the height of the table so that it is usable for all ages and heights. Another issue with the durability of the table was that as a result of the touching surface being made with tracing paper, after extended use the surface could become sticky, dirty, or torn. This issue can cause the table to process “phantom” touches. To counter this problem we changed the touching surface to a more stable Rosco Grey material. This material also helped to block the extra light that the table was processing.

Over half of the participants had previous experiences with using a multitouch device. This experience mostly came from (their parent’s) phones, iPods, and other games. All the participants believed that the table would be useful in classroom. The subject that was selected the most by the children for being most benefited by a multitouch table was math followed by science. It is possible that these subjects were already at the forefront of the children’s minds since they were attending a STEM camp. This could have led to them to not thinking about other subjects such as history, geography, and literature that could also benefit from multitouch table applications. Often during the session, instead of having the children work together to match the animals to the countries, they were told to be competitive and work fast to have the most matches in the first time. This new twist seemed to motivate the children more and resulted in more laughter, faster matches, and more errors. This lack of enthusiasm for collaborating could be why around half of the participants said they would rather work with a partner than alone—perhaps highlighting the need to create better collaborative applications!

Another intervention took place during the ***Harding Avenue Elementary School Computer Club***. This club meets weekly for an hour, and during one of the weekly sessions a 3-person team brought the MET to the session. There were 12 4^{th} and 5^{th} grade students in attendance (7 boys and 5 girls; 10 white and 2 African-American, both females). They had the opportunity to use the MET or to use a game-creation program. To oversee the activity, there was one administrator, two MET developers, and eight undergraduate students.

There was excitement about the MET among students, particularly the boys. The boys showed little hesitation in trying new gestures with the table, grabbing and manipulating the virtual cards in ways not dictated by the rules of the game or activity. For example, the boys were more excited to explore ways in which to disrupt actual gameplay. They would resize a card to fill up the whole board, place objects incorrectly in other boys' bins, or populating the display with an excessive number of cards—laughing enthusiastically while doing so. This would draw attention to the novel technology, but not in a particularly helpful way.

The girls tended to be more hesitant, seeking to learn the rules and to wait for a demo from a facilitator before starting the game. The boys dominated the time with the MET early on, while the girls clustered in one side of the room for much of the session. At first it appeared that they did not like the competiveness of the game that we were playing at the time. After being asked if they wanted to play with the MET, they initially showed disinterest, but later the two African American girls came to the table and took a leadership role in playing the games. It appeared that the girls took the rules of the games very seriously and they wanted to play the games the way they were supposed to be played. During this time the other three girls did not approach the table; it was unclear whether this was because of the number of kids were at the board or because of dynamics with the kids were at the table (the two African American girls particularly seem to aggressively boss around the other children). After the two African American girls and some of the boys left, the other three girls came to the table to see what it was happening. When they were introduced to the table they did not try to play any of the games out right they just played with the cards moving them around and resizing them. They appeared to have no interest in any competiveness that the inherent in the game.

When the students first started to interact with the board, they would only use one finger on the board—as they would with a mouse and cursor icon—even though it was demonstrated to them how to use multiple fingers or hands. They did not start using two hands at a time until they were presented with an action that required to hands. After experiencing the benefits of multitouch, the students used multiple fingers and hands frequently. By the end of the session, all of the students were very social in their interactions, talking and looking at others' activities and games. The MET was a good match for these activities since there were many hands on the table at any time.

With respect to the ***portability and durability*** of the table, we found that disassembly and reassembly of the table usually takes about 15 minutes with one person. The marks placed on the sides, bottom, and top of the table make it easy to assemble the parts in the correct positions. Transporting the table in a vehicle is simple because you can organize the table where the bottom and the sides are stacked on top of each other. The surface top, because it has a glass and sensitive surface, can still on top of the other parts. Because the sides are made of wood, constant adding and removal of the sides can cause the edges of the sides to start to splinter. Another issue is keeping up with all the cables, because the computer, projector, and the surface top all have power cables it can be cumbersome to keep up with all of them. Having a backup computer is essential, in case issues occurred with the computer during travel. Because the table is not completely multitouch, there is also a need to transport a keyboard and mouse as well. An ideal situation would be to have the table that is completely multitouch or at least a cordless keyboard and mouse. There also may be a need to carry a separate extension cord for times where the table cannot be directly near a power source. There were numerous of occasions during demos and school visits where the table was limited to where it could be located because it had to be close to a power source.

6 Conclusions and Future Firections

Multitouch technology supports the manipulation of digital objects using multiple touches—either from a single person or many. This paper explored how our selection of a small number of multitouch gestures representing relationships among objects can support learning applications, toward a toolkit where teachers could create their own digital cards that match their curriculum. Our extension of the card paradigm provides unique advantages over physical cards. By storing the content of the cards in an XML document, the learning activities have been abstracted from the specific educational content. This flexibility supports the use of these cards for teacher-selected content that could be modified by entering topic elements (e.g., animal-country pairs, or animal lifecycle sequences like egg-larva-pupa) and relationship types describing associations between elements.

An area of particular interest for future investigation is in the development of collaborative educational activities for multitouch tables. While many desktop computer applications are inherently single-user, a multitouch table provides opportunities to learn about diverse topics, such as history [6], math [7], music [7], and even basic social skills [8]. Other future work includes quantitative and qualitative study of the multitouch gestures and point and click interfaces, and distribution of applications. Our interventions have demonstrated the promise of multitouch, but a more in-depth exploration is needed. It is our expectation that follow-up work will result in an understanding of how and when technology like multitouch should best be used; e.g., in classroom teaching, as part of after-school programs, or through extracurricular activities like scouting or 4H clubs.

Acknowledgements. Thanks goes to the U.S. National Science Foundation (NSF) for grants IIS-0851774 and CNS-0940358 that supported this work. The opinions in this paper are our own and are not necessarily shared by the NSF.

The authors would like to thank the many people who helped contribute to the construction of the table and to early software instantiations, particularly Goldie Terrell, Stacy Branham, Kristin Whetstone, Chase Carroll and Keith Manville. Thanks also to the many students who used MET, and to the teachers and administrators for their time and insight.

References

1. Buxton, B.: Multi-touch systems that I have known and loved. Downloaded, `http://www.billbuxton.com/multitouchoverview.html` (January 21, 2011) (unpublished manuscript)
2. Han, J.: Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. In: Proceedings of UIST 2005 (2005)
3. Rick, J., Harris, A., Marshall, P., Fleck, R., Yuill, N., Rogers, Y.: Children Designing Together on a Multi-Touch Tabletop: An Analysis of Spatial Orientation and User Interactions. In: Proceedings of IDC (2009)
4. Vallis, K., Williamson, P.: Build Your Own Board. Learning & Leading with Technology (2009)

5. Dorsey, D., George, J., de Araugo, E., Wilson, G., McCrickard, D.S.: Multi-touch tables for collaborative learning. In: Proceedings of the Annual Biomedical Research Conference for Minority Students (2010)
6. Ardito, C., Costabile, M.F., Lanzilotti, R.: Gameplay on a multitouch screen to foster learning about history. In: Proceedings of the International Conference on Advanced Visual Interfaces (2010)
7. Yu, X., Zhang, M., Ren, J., Zhao, H., Zhu, Z.: Experimental development of competitive digital education games on multi-touch screen for young children. In: Proceedings of Entertainment for Education and 5th International Conference on E-learning and games (2010)
8. Piper, A.M., O'Brien, E., Morris, M.R., Winograd, T.: SIDES: A cooperative tabletop computer game for social skills development. In: Proceedings of CSCW (2008)

A Study of Interface Design Method on Crossing Media-Dimension

Hung-Pin Hsu

Graduate Institute of Architecture, National Chiao Tung University,
1001 University Road, Hsinchu, Taiwan 300, ROC
{Hung-Pin.Hsu,ncs}@arch.nctu.edu.tw

Abstract. With the emergence of new types of design media, design products are becoming more diverse. However, the user's habits of designing are difficult to accompany with the design media to convert. This study analyzes the behavior of users to offer an optimized design method for designers using new design media. Therefore, in this paper, furniture catalog is adopted as a study case, for an analysis of used habit of 2D paper-typed catalog. Then, an experiment is setup with a virtual catalog which is designed by 3D interactive software and induced with original used habit. Subjects participate in this experiment and make use of the certain catalog. Finally, two kinds of evaluations, usability evaluation and Heuristic Evaluation, approach to the analysis of new media design. In conclusion, the designer could have a better design method during crossing media-dimension.

Keywords: Human-Computer Interaction, Interface Design, Usability Research.

1 Introduction

With the development of digital technology, it is resulting in the diversity of media forms. This provides designers with more design choices and also to become more diverse variety of design. Moreover, the interaction between user and design also induces to different possibilities. The interactive technology has been used in movies, programs, public art and other designs. Therefore, there have been several research literatures about new type of interactive design [1], [2], [3]. However, most researches only focus on the usability of single design media. The new media and original media are usually analyzed separately. As a result, the relationship between design media and works usability is rarely discussed.

This neglect allows users to produce a lot of cognitive load during the process of carrying out the task operation. For example, the first steps for people undertaking the visual identification is based on the shape and contours of the object. And then, if you want to find the actual location of objects, it mostly needs the perception of depth of visual. That is to say, if the graphic images failed to perform clear and accurate perspective of three-dimensional space, it would lead people to determine a wrong image location or size of objects. Even though spending more effort on the photography and layout, it is still hard to present a complete picture of object [4], [5].

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 247–255, 2011.
© Springer-Verlag Berlin Heidelberg 2011

In order to improve such flaws, designers begin to design products through interactive technology. Nevertheless, they do not take original used habits into account in the design process, so that the new design work made by new design media is like as a whole new product but just similar to one made by original design media. In other words, there is no close connection about used habit between new and original design work.

Therefore, the purpose of this study is to find out a method that could smoothly continue the original used habits into new design work no matter which is made by new or old design media. In addition, this study takes the furniture catalog as a study case, which features user task and visual dimension. And moreover, the following research question is that: what are the habits of users when they are in the use of flat paper catalog? How do designers continue the original used habits into new design media product? And what will be the design evaluation based on the original used habits?

2 Methodology and Steps

In order to solve the previous research problems, there are two steps in this research. First, this study analyzes the usability behavior with the original media. Second, these analyzed data were taken to transfer to the new media design. Then, the evaluation of usability on new media design was also been analyzed by user and expert.

2.1 The Analysis of User Behavior on the 2D Design Media

In this step, this study takes 2D catalog to be the research object represented the original design media. There are thirty users of 2D catalog who have been invited to join this experiment. In order to get more plentiful data, this study also interviews five stakeholders who work in the furniture store with the supply of 2D catalogs.

Issue Survey and Analysis. To get the specific tasks that user could encounter in the use of 2D catalog, this study executes issue survey before the user task analysis. In this stage, this study invites thirty users and five stakeholders to describe and demonstrate the usage of 2D catalog. After collecting the interview data, this study constructs an affinity diagram to find out the main targets of users. Then the user's goals could be found by generalizing these targets. The user's goals and targets when they are using 2D catalog is shown in the Table 1.

In the Table 1, the user's goal of category A is that user tries to understand the characteristics of products. In the target A1, user tries to understand the scale, form, color, material, designer and the design notion of products. And in the target A2, user also tries to understand the price, weight, capacity and place of origin. The user's goal of category B is that user wants to learn how to know the relationship between products and themselves. It includes that how to choose a product suitable for faculty of body in target B1, for example, choosing a pillow suitable for the height of neck. The target B2 describes that user wants to learn how to choose a product suitable for faculty of life, for example, choosing a multiple function office furniture to improve the working efficiency. The target B3 is that user wants to learn how to use and fabricate the product; it is about the method of fabrication, clean, reposition and maintenance. The user's goal of category C is that user tries to get the information about collocation. The information in target C1 is about collocating product with other products, such as collocation between a desk and a lamp. And the target C2 is the information about collocation between product and space, such as the indoor planning, design style, light design, scenario simulation, etc.

Table 1. The user's goals and targets

User's Goals	User's Targets
A. To understand and compare the information of products	A1. To understand and compare the form of products A2. To understand and compare the function of products
B. To learn the knowledge of choosing and using products	B1. To learn how to choose a product suitable for faculty of body B2. To learn how to choose a product suitable for faculty of life B3. To learn how to use and fabricate the product
C. To get the information about collocation	C1. To get the information about collocation between products C2. To get the information about collocation between product and space

Besides user's goals mentioned above, we could find out during interview that users expect to get other information in the use of 2D catalog. This includes spirit of enterprises, life attitude, the way of paying, after-sale service, information and activity of store, etc. In conclusion, this study keeps these requests of users to make sure that the new media design satisfies users.

User-task Experiment with 2D catalog. This stage is to execute the user-task experiment with 2D catalog. Each of thirty subjects, the same as ones in issue survey, has forty-five minutes to execute the user's task in the experiment environment. This experiment supplies subjects with a 2D catalog, a spatial layout and a scenario essay which is compiled with the user's targets from issue survey. The scenario essay is as Table 2.

Table 2. The scenario in the user-task experiment

You Got an Opportunity to Reconfigure a room FOR FREE!!
You could pick up something from catalog such as a lamp you like, a pillow suitable for your neck or a desk improve working efficiency and even a cabinet to collocate the lamp. All these furniture would be free as long as you help us to describe the form and function of the lamp, how you choose the pillow and desk, as well as the way you collocate the cabinet and how to fabricate it. Now start to choose your furniture for your room!!

After executing user's tasks, the subjects proceed with an interview based on their situation in the experiment so that this study could acquire more complete information about subjects in the use of 2D catalog. Moreover, this study would set up one camera to record the facial expressions of subjects and the other to record the interaction between their hands and 2D catalog. In addition, a directional microphone is also set up to record the verbal data when subjects use the think-aloud method to execute the task during experiment.

Construction of User's Behavior Model. To analysis the interaction data, Contextual Inquiry methodology is chosen to build the cognitive and multi-personal interaction models [6], [7]. Five kinds of models are built based on this methodology: Flow Models show the relationship between roles. Cultural Models show the influence of culture in the interactive behaviors. Sequence Models show the targets and steps that users execute in the interactive activities. Artifact Models show the artifacts that users use in activities, such as tools, objects, products, etc. Physical Models present the environment of the interaction.

Since this study is to identify the used habit in the process of user tasks, these used habits would be regarded as kinds of design factor and be adopted in the process of designing 3D catalog. Moreover, the users are the same no matter in new catalog or original catalog. So models of 3D catalog, those are flow, cultural, artifact, and physical, would be also the same as those of 2D catalog. According to these principles, this study takes the behavior data obtained from user-task experiment with 2D catalog to construct the sequence model which could describe the user's goals and steps. The sequence model of 2D catalog is shown as Figure 1.

At the course of the experiment, we would observe that much of information that user cares about is not presented in the 2D catalog. For example, although there are explanations of words and figures about the size of products, subject still feels not complete. This is because that subject looks at multiple products in the same page, and these different sizes of products are shown in the same scale of figure so that it would make subject misunderstood the correctness of real size of product. If the figures of products are overlapped with each other, it would reduce the opportunity to present the appearance of products.

In addition, for those of subjects who have few concepts about spatial scale, it is hard to understand the correct scale through the text. The subjects also worry about the part not shown in the figure might be different with their expectation, the backside of product for example. Even the color and texture of products might be discrepant as a result of photo-shooting skill and printing. Subjects also indicate that to put price tags beside the products could always attract their attention for they could immediately determine to buy or not according to their financial capacity. Furthermore, a price tag beside the product represents that this product is now on sale even though multiple products are presented in the same figure.

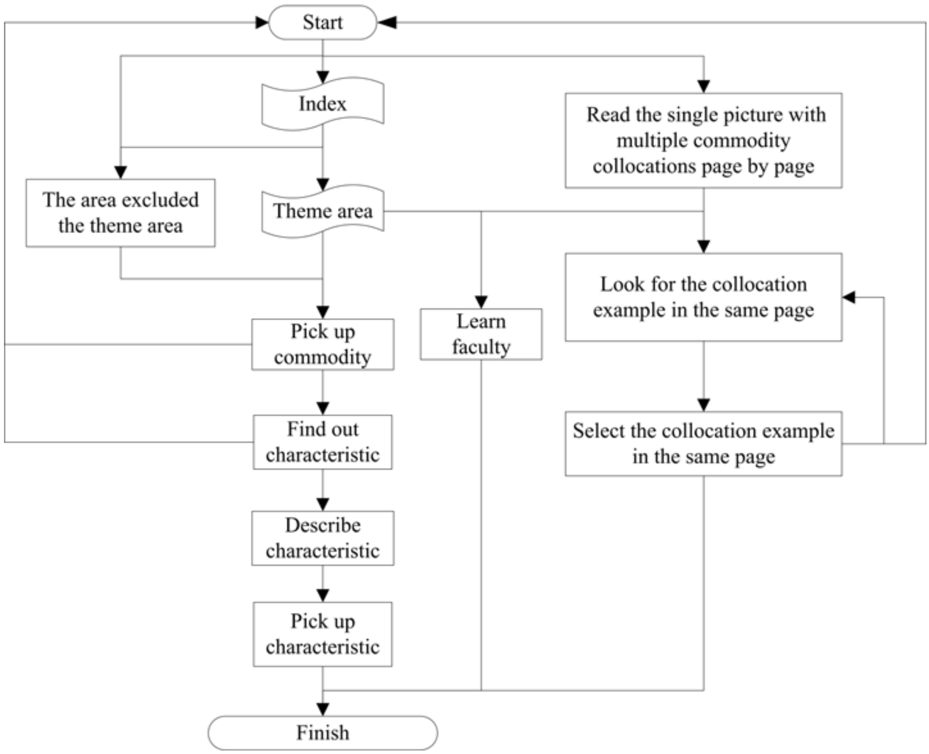


Fig. 1. The sequence model of 2D catalog

2.2 Design and Evaluation of 3D Design Media

In the beginning, this study undertakes to design 3D catalog with the behavior model and date of 2D catalog. Then we invite 30 subjects the same as ones in 2D catalog to execute the user testing of 3D catalog. In addition, there are 6 experts in interaction interface design invited to execute the heuristic evaluation on 3D catalog. Finally the opinions from users and experts are integrated. [8]

Design Method of New Type Media. This study uses AutoDesk 3ds Max and Quest3D as 3D interaction design tool. The content of 3D catalog is a 3D interactive virtual scene, where users could look around the 3D product model. The scene also provides voice commentaries for users to understand the information of product. Meanwhile, the operation way is to control the rocker forward and back. The left bottom is designed to open and close the menu, and the right bottom is to switch the initial rotation setting of XY-axis plane to YZ or XZ-axis, for viewing in 360 degrees.

According to the result of user-task analysis of 2D catalog, this study proposes three main points of the new media design. First point is to provide a product in first person viewpoint to exhibit three-dimensional space of 3D catalog. This metaphor of design is based on the 2D catalog and the physical store in real world. There are two index models, spatial and product categories, for users to browse. For example, to browse a lamp with a theme of living room, or present all the lamps at the exhibition center. Before beginning, users have to set their height so that they can tour the space at correct height. Besides these two kinds of index mode, user could also choose a specific product to view in 360 degrees.

Second, this study provides a menu for choosing the functions. The menu include index of product, index of marking, index of space classification, height marking, picture marking, teaching information area, voice explanation switch, 360 degrees viewing and route recording. In particular, the picture marking is designed for marking the important page and the metaphor is set as camera photography due to the first person viewpoint.

In third, this study designs a dynamic trigger method for showing the information. When user closes to the products, the dynamic trigger would present the information of the product. And the information would hide automatically when user leaves from the product. This design concept could reduce the information load efficiently. Moreover, user could also use the 360 degrees viewing function to regulate the environmental light and control the function of product.

Evaluation of 3D Catalog. In this stage, subjects are invited to execute the user testing with 3D catalog. The experiment environment, device and task are the same as the 2D catalog experiment. The reason why we invited the same subjects is to let them know their suggestions have already been considered in the design process of 3D catalog. To compare the difference of operating steps between 2D and 3D catalog, this study uses the data from user testing of 3D catalog to build the sequence model. The sequence model of 3D catalog is shown in the Figure 2.

In order to get more complete evaluation data, this study also invites five experts to use the Heuristic Evaluation methodology to offer some advice for 3D catalog design. The "Ten Usability Heuristics" [9] is chosen in the heuristic evaluation process. To let experts could compare the difference between 2D and 3D catalogs; this study also provides 2D catalog and the sequence models of 2D and 3D catalogs. Finally, the user's suggestion from user testing is combined with the result of heuristic evaluation. The evaluation results of using 3D catalog are shown as below:

1stUsability Heuristics: Visibility of system status. The dynamic trigger presents information, name and price of product. It would let user easily know which product they could buy. As the menu is unfolded, there would appear a small window in the place where the cursor is moving on. And user could easily know the meaning of function and decide to choose it or not.

2nd Usability Heuristics: Match between system and the real world. The metaphor of 3D catalog that user walks in the store are the same as the real world. Users usually bow their heads to read the menu, but the way that menu unfolds in the 3D catalog seems as someone put a paper in front of their face. Also, before users view in 360 degree, they need to set up the rotate parameter of the screen. However, this request is difficult for the user without an axial concept.

3rd Usability Heuristics: User control and freedom. User could walk around the 3D scene through controlling the rocker, but the system cannot allow the user to sit or lie down in bed. So it is difficult for user to choose a product suitable for the faculty of body.

4th Usability Heuristics: Consistency and standards. Because the design style and words in the 3D catalog follow the standard in the 2D catalog, users who used to use 2D catalog would not feel unfamiliar when they contact with the 3D catalog in the first time.

5th Usability Heuristics: Error prevention. Because users could connect system functions through the menu and record routes so that users could use the menu to search their location when they lose.

6th Usability Heuristics: Recognition rather than recall. 3D catalog provide users with remark function to memorize the important location and scant their cognitive load.

7th Usability Heuristics: Flexibility and efficiency of use. After using route-recording function in the menu for a while, the degree of dependence that user use it to remark and search passed location is raising. This function helps users improve work efficiency with the increase of experience.

8th Usability Heuristics: Aesthetic and minimalist design. There is only one menu on the operating interface, and the option levels of menu are not too many, so that users could choose the functions easily. In the other hand, some design issues have to be considered. For example, as users view the screen in the 360 degree, if it is suddenly jumping to another scene, this design point concerns users about how to come back to the original scene. This is because the continuity does not be included into design consideration. Unfortunately, the same situation also occurs in switching the browse modes. In conclusion, the suggestion is that interface designer must try to handle the transition or perform different functions at the same screen as much as possible.

9th Usability Heuristics: Help users recognize, diagnose, and recover from errors. There are few users losing their way in the virtual environment. Meanwhile, most of them would open the menu to check the location where they are. This design point help users find their way to go.

10th Usability Heuristics: Help and documentation. Information signboards are set up in front of the entry of store and each theme area. These signboards could be read directly and trigger the voice of instruction when users are close to them. However, most of users would ignore the signboard at the entry, and the other users who read the signboard would not spend more time listening whole the instruction. In contrast, most of users pay more attention to the signboards set at the entry of theme areas.

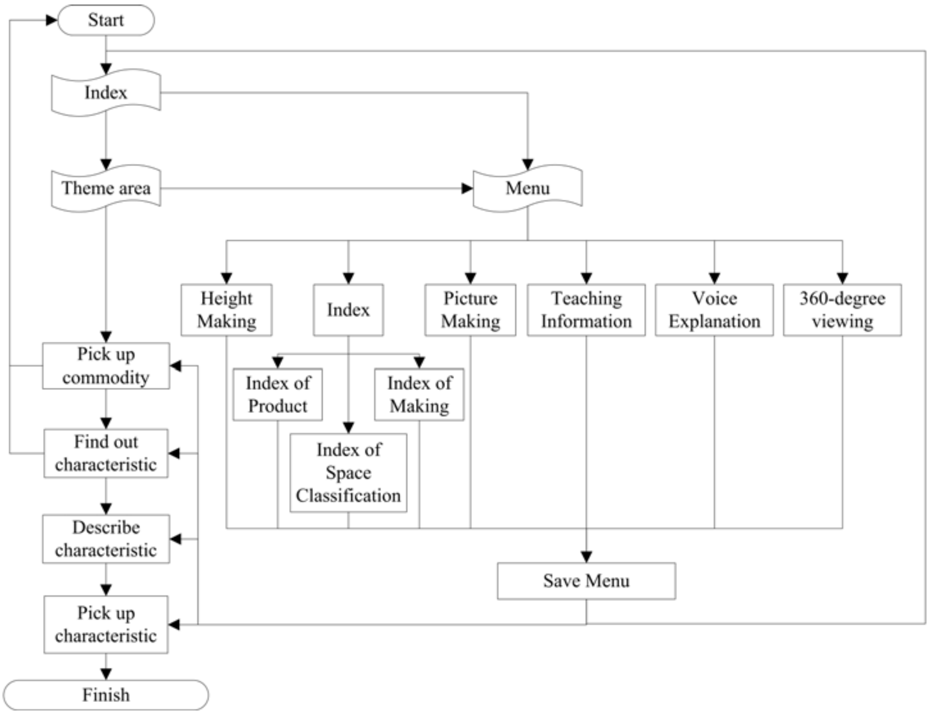


Fig. 2. The sequence model of 3D catalog

3 Conclusions

This study has already presented a 3D interactive catalog as a case designed by new design media. The interactive scene could let user intuitively view, study and collocate products in the metaphor of walking in the store. Through the methodology of this study, designer could cross the limitation on the dimension of design media, and moreover, to transfer used habit from 2D to 3D media design. According to the evaluation result of users and experts, the design of 3D catalog has already been improved and reduced the defect of 2D catalog. Meanwhile, there is no new problem on usability in the 3D catalog. More importantly, this methodology could provide designers in different field to transfer the used habits of users to the new design. However, because of resource constraints, this study only takes the catalog of furniture as the sample. Despite the reason, the methodology of this study could be verified with other design research in different fields.

References

1. Crilly, N., Moultrie, J., Clarkson, P.J.: Shaping Things: Intended Consumer Response and the Other Determinants of Product Form. Design Studies 30(3), 224–254 (2009)
2. Chamorro, K.M., Popovic, V., Emmison, M.: Using Visual Representation of Concepts to Explore Users and Designers' Concepts of Everyday Products. Design Studies 29(2), 142–159 (2008)

3. Solso, R.L.: Cognition and the Visual Arts. MIT Press, Cambridge (1996)
4. Hackos, J.T., Redish, J.C.: User and Task Analysis for Interface Design. Wiley, New York (1998)
5. Kuniavsky, M.: Observing the User Experience: A Practitioner's Guide to User Research. Morgan Kaufmann, San Francisco (2003)
6. Beyer, H., Holtzblatt, K.: Contextual Design: Defining Customer-Centered Systems. Morgan Kaufmann, San Francisco (1997)
7. Preece, J.A.: Guide to Usability: Human Factors in Computing. Addison Wesley, Wokingham (1998)
8. Shneiderman, B.: Designing the User Interface. Addison Wesley, Wokingham (1997)
9. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods. John Wiley & Sons, New York (1994)

Multi-language Online Word Processor for Learners and the Visually Impaired

Shiblee Imtiaz Hasan

Department of Computer Science and Engineering, BRAC University,
66, Mohakhali C/A, Dhaka 1212, Bangladesh
shiblee@ieee.org

Abstract. Most of the modern operating systems come with accessibility and screen reading features. They also have with basic word processing applications. However, many of the languages are still not supported in the screen reading application and neither do they come with features such as translation and transliteration (phonetic). Many plugins are available to overcome these barriers, but the visually impaired users and/or users without administrative privilege will not be able to install those in the local computer. This paper discusses about implementation of a rich internet application that enables users to have access to a free online word processor which can translate, transliterate and speak out words (in different languages) that have been typed, which can be very helpful for learners of foreign languages and visually impaired users.

Keywords: languages, word processor, speech synthesis, transliteration, accessibility.

1 Introduction

Desktop-based applications, especially word processors are usually known to be more powerful than the online ones. They also come with more features compared to the Internet applications. However, the online applications can be very portable and can run without being installed in a computer. This paper would discuss about an online word processor which can be portable solution for many. The application can be a very important document processing utility for the visually impaired too. Some of the features would also assist learners of a new language.

There are different word processing software available in the market. Some are commercial e.g. Microsoft's Word or Apple's Pages, while others are open-source e.g. OpenOffice.org. There are also some word processors that are web browser-based Rich Internet Applications (RIA) such as Google Docs and Zoho Writer.

2 Features Overview

The application contains features such as automatic transliteration and translation in different languages, text-to-speech (or speech synthesis), text formatting, download or email document, keyboard controls etc. The transliteration is used for languages other

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 256–260, 2011.
© Springer-Verlag Berlin Heidelberg 2011

than English, so that those who are familiar with the English keyboard can easily get the right word in a different language. This would also allow learners to know the right spelling of a word without manually checking the dictionary.

The text-to-speech speaks out the word that has been typed recently which enables the visually impaired to know if he/she typed the right word. For learners, it is also a good tool to learn the pronunciation of any specific word. The text formatting permits the users to create formal documents with proper text styles and fonts. The document is automatically saved in a user's account if the person is logged in or it is saved in the server session for that specific computer. So even if the browser crashes, the data will be saved in the server for the user to retrieve.

If the user wants to share the document, it could either be downloaded or sent by email from within the application in different formats. All of the functionalities can be accessed by using combination of keyboard keys, which is very essential for the visually impaired users.

3 System Architecture

The interface of the application has been built from Hypertext Markup Language (HTML), JavaScript and Asynchronous JavaScript and XML (AJAX). In the backend, PHP conducts operations such as communication between interface and text-to-speech engine, speech recognition engine; saving the document state etc.

The speech engines have been installed in a Linux server. The English text-to-speech is called Festival, developed by University of Edinburgh and the Bangla voice of Festival is developed by Center for Research on Bangla Language Processing (CRBLP), BRAC University. The translation and transliteration features will make use of the Google API, which contains a large database of words and meanings in different languages.

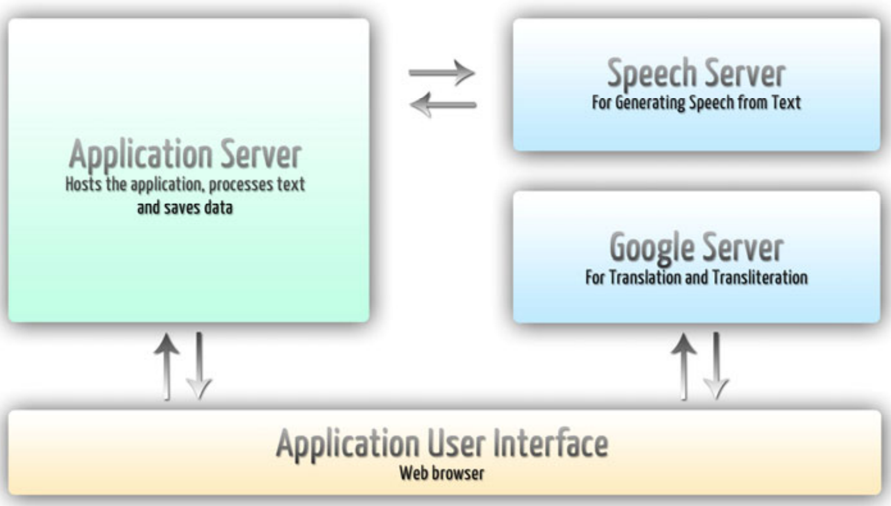


Fig. 1. The system architecture

The JavaScript interface communicates with both the application server and Google's server. This enables a faster translation and transliteration service because the data doesn't go through the application server.

4 User Interface

As mentioned earlier, the Graphical User Interface (GUI) is created with HTML and JavaScript. The interface is designed with shades of gray. This can be helpful for the color-blind persons. There is a large menu button that can assist the partially visual impaired persons. The menu button can adjust the colors as well as the size of the buttons. The features that are enabled are shown in invert color, which makes a clear contrast with those who are not in use. The features can be accessed by keyboard combinations so that visually impaired can use them.

Fig. 2. Screenshot of the word processor, showing the basic functionalities and the transliteration enabled (*button in black background*)

5 Speech Synthesis

For speech synthesis or text-to-speech feature, the Festival Speech Synthesis system has been used which is open source and runs in a Linux platform. [4]

Festival offers a general framework for building speech synthesis systems as well as including examples of various modules. As a whole it offers full text to speech through a number APIs: from shell level, though a Scheme command interpreter, as a C++ library, from Java, and an Emacs interface. Festival is multi-lingual (currently English and Spanish). Due to being open-source, many other groups release new languages for the system including the Bangla voice by CRBLP.

6 Translation and Transliteration

For machine translation and phonetic transliteration, Google's services were used. Google Translate is a statistical machine translation service provided by Google Inc. to translate a section of text, document or webpage, into another language. This can be used freely by using the APIs provided by Google.

Google Translate, like other automatic translation tools, has its limitations. While it can help the reader to understand the general content of a foreign language text, it does not always deliver accurate translations. Some languages produce better results than others. However, because of being web-based, it is easier to implement it.

Google transliteration is a transliteration typing service for languages. This tool from Google is based on dictionary based phonetic transliteration approach. In contrast to older Indic typing tools (which type by transliterating under a particular scheme), it transliterates by matching the Latin words with an inbuilt dictionary. So, users do not need to remember the transliteration scheme due to which this service is very easy and suitable for first time typists or beginners.

7 Unicode and Complex Scripts Support

Even though many operating systems are able to display Unicode characters, most are unable to display complex glyphs if some specific fonts or programs are not installed. To overcome the missing font issue any of the following methods can be used.

7.1 Cufón

This is a Javascript library which can be used replace text with canvas and VML. This allows the browser to display fonts which are not installed in the system. Cufón consists of two individual parts – a font generator, which converts fonts to a proprietary format and a rendering engine written in JavaScript.

7.2 sIFR

The Scalable Inman Flash Replacement or sIFR is an open source JavaScript and Adobe Flash dynamic web fonts implementation, enabling the replacement of text elements on HTML web pages with Flash equivalents. It is a scalable variety of HTML text-to-flash replacement. Despite of being inside Flash, accessibility can be preserved using this technique.

Despite of having necessary fonts in the system, the system might not be able to render the fonts for complex scripts if there is no native support in the operating system. To overcome this issue, Adobe Flash can be used to properly render the glyphs of complex scripts.

Text Layout Framework for Adobe Flash Player 10 and Adobe AIR 1.5 is an extensible library, built on the new text engine in Adobe Flash Player 10, which delivers advanced, easy-to-integrate typographic and text layout features for rich, sophisticated and innovative typography on the web. It supports bidirectional text, vertical text and over 30 writing systems including Arabic, Hebrew, Chinese, Japanese, Korean, Thai, Lao, the major writing systems of Bangladesh and India, and others.

8 Future Developments

Several more features have been planned for implementation in the system. They include file import/export, online spelling and grammar checking with speech synthesis etc.

References

1. Alam, F., Nath, P.K., Khan, M.: Text To Speech for Bangla Language using Festival. In: Proc. of 1st International Conference on Digital Communications and Computer Applications (DCCA 2007), Irbid, Jordan (2007)
2. Clark, R.A.J., Richmond, K., King, S.: Festival 2 - build your own general purpose unit selection speech synthesiser. In: Proc. 5th ISCA workshop on speech synthesis (2004)
3. Taylor, P.A., Black, A., Caley, R.: The architecture of the festival speech synthesis system. In: The Third ESCA Workshop in Speech Synthesis, Jenolan Caves, Australia, pp. 147–151 (1998)
4. Black., A.W., Taylor, P.A.: The Festival Speech Synthesis System: System documentation. Technical Report HCRC/TR-83, Human Communciation Research Centre, University of Edinburgh, Scotland, UK (1997), `http://www.cstr.ed.ac.uk/projects/festival.html`

A Guideline for an Outpatient Guidance System for Use in General Hospitals

Gunhee Kim[1], Jukyung Park[1], Manchul Han[1], Hyunchul Park[1], Dayoung Jung[1], Sehyung Park[1], Seongil Lee[2], and Sungdo Ha[1]

[1] Intelligence and Interaction Center,
Korea Institute of Science and Technology, Seoul, Korea
[2] Department of Systems Management Engineering,
Sungkyunkwan University, Suwon, Korea
{kani,parkjk,manchul.han,hcpark,seria26,
sehyung,s.ha}@kist.re.kr, silee@yurim.skku.ac.kr

Abstract. This paper presents a guideline for a hospital guidance system that provides outpatients with guidance about their process and location. When visiting general hospitals, outpatients often encounter difficulties in understanding the hospital process and in knowing where to go next during that process. To address this issue, we initially conducted an ethnographic interview of fifteen outpatients to determine their needs, interviewed hospital staff to discover exceptional cases, and then analyzed the hospital process for outpatients to uncover any timing difficulties so as to specify the proposed guidance service. Considering the findings of the user study and the result of a domain analysis, we formulated the outpatient guidance service. The outpatient guidance system is presented here as an example.

Keywords: Outpatient guidance system, General hospital, Outpatient services, Process, Location.

1 Introduction

When visiting public hospitals, outpatients often experience difficulties in understanding where to go and what to do. Although hospital personnel are ready to provide necessary guidance, they cannot always address the difficulties of all outpatients because these issues are very diverse.

Context-aware computing has appeared as a viable means of providing proper guidance to users in difficulty. In an attempt to provide guidance based on a user's context, researchers who study Location-based Service (LBS) have suggested several systems that give information relevant to a user's current position [1, 2]. Location-based information can be helpful to users in these circumstances, but these systems are limited in terms of their ability to provide more customized guidance to a user requiring a personalized process. Han [3] presented a spatiotemporal context-aware system that guides users considering both the user's location and process, but it is targeted for public places; therefore, it is too general to meet the expectations of outpatients in general hospitals.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 261–269, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Several context-aware guidance systems in hospitals have been developed. A location-aware hospital system [4] was implemented to give patient information to hospital workers, and a context-aware nurse call system [5] was designed. However, these systems are designed for hospital staff only. For outpatients, a patient guidance system [6] was suggested which gives guidance about a user's current step in the hospital process. It can provide them with their waiting number on electronic paper. However, this patient guidance system only considers a restricted set of process conditions; hence, its functions are too simple to cover the wide variety of needs patients typically have.

In this paper, we present a guideline for an outpatient guidance system for general hospitals. To understand a user's difficulties in general hospitals, we used relied on observations and interviews and then analyzed the process used by outpatient services. With the results, we defined the necessary guidance services for outpatients.

2 User Study

A user study was conducted to determine the difficulties encountered by people when they visiting a general hospital as an outpatient. To gather this information, we relied on ethnographic interviews [7] that included user observations and general interviews. In order to determine the proper outpatient services, we conducted direct observations and interviews of outpatients. Subsequently, subject-matter expert interviews were done to discover any unusual or abnormal situations that were not perceptible in the observations.

2.1 Ethnographic Interview: User Observations and Interviews

Setting. Observations were done in a general hospital in an effort to determine the types of difficulties. This method also helped to define the types of guidance that should be provided to users. The observations were performed using an ethnographic method of the type typically used to evaluate the behavioral patterns of subjects. To establish the difficulties encountered by outpatients, the observation procedure was performed at *Seoul National University Bundang Hospital*, which is one of the leading hospitals in Korea due to its state-of-the-art digital infrastructure. There are 3,600 outpatients who visit to the hospital every day. However, there are only about 35 guides to assist them, implying a workload of approximately 100 outpatients per guide per day.

Participants. We observed five male and ten female outpatients who used the outpatient services of the hospital. The average age of the outpatients was 56, and they had visited this hospital 2.8 times on average.

Method. The observation proceeded throughout the overall outpatient experience, from reception to their last service, such as making a payment. The observations continued for 120 minutes on average. The observation group consisted of two people: a helper who is an expert in outpatient services, and an observer who transcribed all of the difficulties experienced by the subjects while observing from the sidelines. In particular, we decided that the helper should not provide guidance to subjects proactively, but guide passively only when the subject had a question that required an answer. We expected that this method, assigning one helper, could help to discover more difficulties than those noted only from observations at a distance, as subjects were believed to express their difficulties only rarely. First, we obtained the consent of the outpatients that we intended to observe. We then instructed the outpatients to tell the helper immediately whenever they required help, needed support or wanted to know anything. While performing the observation, whenever the subjects encountered any difficulty, the helper assisted subjects according to their requests and the observer recorded the particular type of trouble, such as how the helper aided the subject, using a checklist that included the location, time, difficulties, and the comments of the helper. Additionally, subject interviews were performed midstream and after the observations. These were performed while waiting for the consultations or examinations, as the subjects were not always mindful of their difficulties once they had been resolved.

Findings. After the user observations and interviews, we gained an understanding of the typical user type and their difficulties. We also uncovered a number of similarities among the difficulties experienced by outpatients. Every outpatient wished to know about the process and location of outpatient services. 125 similar difficulties gathered from the result of the observations and interviews were classified into two major groups: Process-related and Location-related difficulties. To define the required guidance for outpatients more specifically, the Process-related difficulties, which represented the majority, was classified into four minor groups: Procedural, Time-related, Method-related and Content-related (see Table 1).

1. Procedural: Questions such as “What do I do next?” and “What is my entire process?” are included in this category. Thus, difficulties related to the process were classified into this category. 18 (14.4%) of the difficulties were classified into this category.
2. Time-related: Difficulties related to the waiting time, last visiting date, next visiting date and appointment time are classified in this category. 22 (17.6%) of the difficulties were classified into this category.
3. Method-related: This category contains difficulties in using equipment or performing a task, such as how to use a payment machine or whether a patient should take off his/her shirt for an X-ray examination for instance. 24 (19.2%) of the difficulties were classified into this category.
4. Content-related: In many cases, the observations revealed that outpatients wanted more information in greater detail about their examinations or illnesses. Occasionally, they were also unclear about how much they had to pay. 18 (14.4%) of the difficulties were classified into this category.
5. Location-related: As is the case with people in many complex places, it was found that the outpatients in this study typically had difficulty finding their way around the hospital. Furthermore, it was found that they commonly experienced some difficulties with this even when maps were provided. 43 (34.4%) of the difficulties were classified into this category.

Table 1. Five categories of difficulties

CATEGORY	Procedural	Time-related	Method-related	Content-related	Location-related
Numbers of cases (%)	18 (14.4)	22 (17.6)	24 (19.2)	18 (14.4)	43 (34.4)

2.2 Subject-Matter Interviews: Hospital Staff Interview

Settings. Subject-matter expert interviews were done to discover any unusual or abnormal situations that were not perceptible in the observations. Seven faculty members who were interviewed had worked at various stations of *Seoul National University Bundang Hospital* for 11.8 years on average as of the date of this study. One from the Joint Disease & Reconstruction Center, two from the examination department, one from the reservation desk, one from the reception desk, and two from the information desk.

Findings. From the interviews of the hospital staff members, we discovered exceptional situations that were not found during the observations. Examples include a guardian's information handover problem, a schedule change problem, missed appointments, a different person with the same name, and a fasting-related problem. These problems should be also covered by a guidance service, as should the difficulties noted from the user observations.

3 Hospital Domain Analysis for Outpatient

To determine the timing of the difficulties and specify guidance services, we first analyzed the hospital process for outpatient services. This analysis was conducted in conjunction with a medical information team at *Seoul National University Bundang Hospital*. Figure 1 shows a flow chart of the outpatient services offered at this hospital. Generally, outpatients have at most nine service steps to complete their goal, and the service steps can be reduced according to their situation. To finish each service, users should perform an average of four tasks.

Movement is the most common task that outpatients have to do. Whenever services are finished, outpatients should find their way to the next service. Moreover, when outpatients want to go a specific place that is not an outpatient service place, it is more difficult for them to find their way by themselves when they have not received any guidance from the staff. Therefore, location guidance is the most essential guidance service for outpatients. Second, guidance to the next service and task is necessary for outpatients because there are many prioritized services and tasks that patient should perform, and the priority can change in special cases. In addition, the current day's schedule should be provided because the outpatient service procedure can be long and varied after the initial consultation. Third, it is important to provide information related to waiting and reservation times, as outpatients spend most of their time waiting for their turn. Finally, a sufficient explanation of an outpatient's current service is necessary. Particularly, when they use the payment and examination services, we found that many outpatients do not know what they are paying for or why they must take some tests.

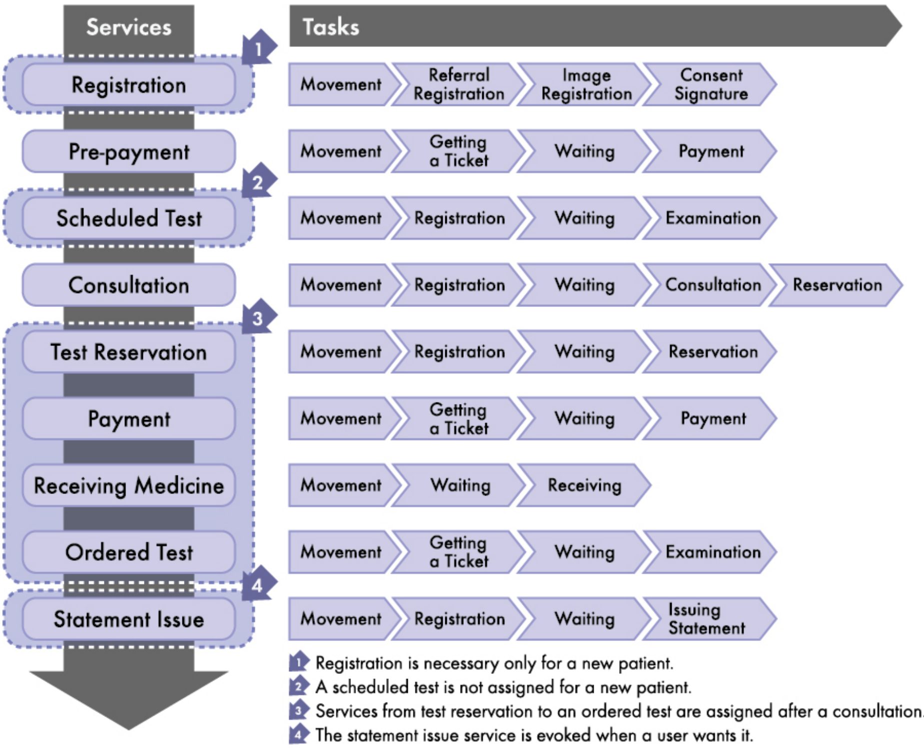


Fig. 1. Flow chart of outpatient services

4 Outpatient Guidance Services

Based on the five categories of difficulties and the result of the hospital domain analysis, we designed hospital guidance services for outpatients (see Figure 2). Guidance services for the procedural, time-related, method-related, and content-related difficulties are grouped into process guidance that includes task, notification, schedule and detailed information guidance. In addition, location guidance is defined for the location-related difficulties related to the next destination and guidance to a specific place.

Process Guidance. The goal of process guidance is to enable outpatients to understand the current services and to know precisely what to do. Process guidance includes task guidance and notification guidance as primary types of guidance, as these are essential for outpatients to follow the hospital procedure. Moreover, schedule guidance and detailed information guidance are given as extra types of guidance that outpatients may want.

Task guidance provides information about an outpatient's current service and guides the user to complete the current task so that the outpatient can understand where he/she is in the process and know what to do at that time. Figure 3a shows an example of task guidance in a test service. As regards task guidance, outpatient guidance systems should be able to understand a user's current service and the task steps in the service.

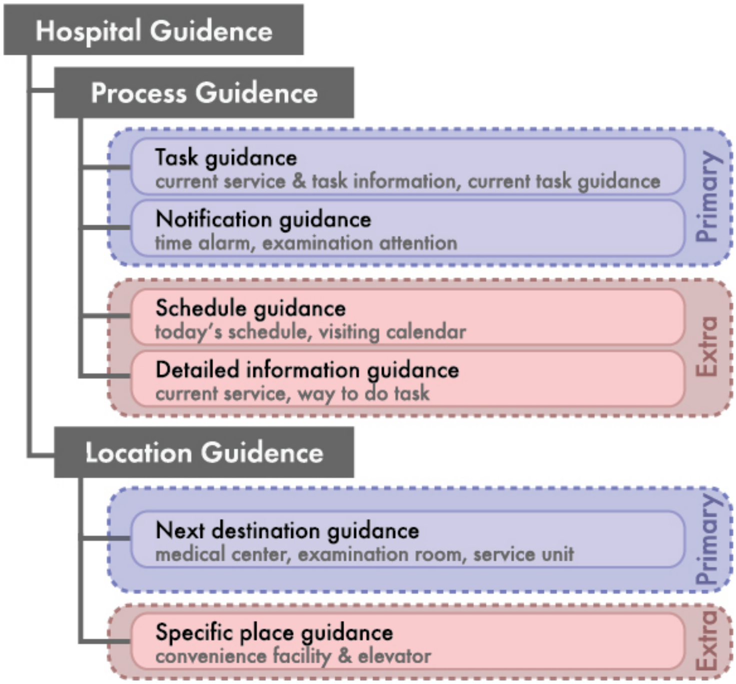


Fig. 2. Hierarchy of outpatient guidance services

Notification guidance proactively sends important messages in the form of alarms that outpatients should recognize. This notification prevents outpatients from missing their turns or violating instructions such as fasting before an examination. For example, when user is called for his turn but the user is not in the waiting room, a notification message with an alarm, such as vibration, is delivered. For notification guidance, guidance systems should monitor the user's context-related time or examination time.

Schedule guidance informs patients of the current day's procedures and hospital visit date, including past and/or reserved appointments. Outpatients can check their remaining services and anticipate the time they will need with schedule guidance.

Detailed information guidance gives more detailed information about the current service or shows patients how to do tasks. Outpatients can obtain more information about their services, such as information on his/her doctor, test process, or payment items. With this information, outpatients can understand why the services are necessary.

Location Guidance. Location guidance is essential because outpatients should move to the next service place whenever a service has finished or when they are looking for a specific place, such as a toilet. Although there are many signs and maps in hospitals, it can be difficult for outpatients to find the necessary signs or to understand guide maps. Therefore, two types of location guidance are required: overall location guidance such as maps, and detailed location guidance such as signs.

Next, destination guidance provides a path from the current position to the user's destination. When a service is finished and a new service is set, guidance systems should be able to recognize a new destination automatically and start to provide guidance to the next destination.

Specific place guidance is not primary for hospital services, but this guidance is useful because many of the outpatients in our user study asked about the locations of convenience facilities and elevators. In particular, there is much waiting time in hospitals, and outpatients often want a specific place to go to wait.

5 Example: A Hospital Guidance System

To describe a more tangible guideline, we illustrate a hospital guidance system using a mobile phone with two guidance modes: process and location guidance. According to the guidance hierarchy, the process guidance mode describes guidance information about Schedule, Task, Notification, and Detailed information guidance. The default page of the process mode gives current-service information and notification messages such as the waiting time (see Figure 3a), and the overall process page shows the current day's procedures along with the current step. Furthermore, patients can check the schedule for a reserved examination or for other information (see Figure 3b). The location guidance mode was also designed to have two pages: the current view and hospital map pages (see Figure 4). The current view page displays a picture with an arrow that is used to inform a patient as to whether they should turn or go forward; the hospital map page shows the overall route for the next destination or specific places such as toilets.

The user interface of a hospital guidance system should be designed for the elderly, because a considerable number of patients are older people; moreover, the size of the text should be large enough for the elderly to read it, and the method of changing the page should be very simple.

(a) Current service page (b) Overall process page

Fig. 3. Pages of process guidance mode

Fig. 4. Pages of location guidance mode

6 Conclusion and Future Works

We introduced a guideline for an outpatient guidance system that provides outpatients with guidance about their process and location. With the increased popularity of smart devices such as smart phones and tablets, the need for personalized guidance in complex environments has increased. Particularly, in general hospitals, outpatients should pay attention and follow passively the directions of hospital staff owing to the current shortage of human guides or guidance systems. As the situation currently stands, outpatients pay occasionally without any information about the payment items. We believe that the proposed outpatient services will be helpful to users and that applying this guideline can make hospital guidance systems more intelligent and effective.

To reinforce our guideline, first we are planning to conduct long-term user evaluations in the near future. We expect that these evaluations, with mobile devices, will help uncover new difficulties and solutions. We also plan to research context-aware methods to understand users' situations in hospitals, as some difficulties cannot be covered due to the limitations associated with context-awareness; for example, some services are not updated in real time regardless of whether or not they are finished, causing the system to infer with the situation by combining the user's locations. We hope to overcome these limitations in future research.

Acknowledgments. We specially thank to Seoul National University Bundang Hospital for helping to conduct the user study and domain analysis. This work was supported by the IT R&D Program of MKE/KEIT [KI001824, Development of Digital Guardian technology for the disabled and the aged person].

References

1. Stephanidis, C., Paramythis, A., Zarikas, V., Savidis, A.: The PALIO framework for adaptive information services. In: Multiple User Interfaces: Cross-Platform Applications and Context-Aware Interfaces, pp. 69–92. John Wiley & Sons, Chichester (2004)
2. Paganelli, F., Bianchi, G., Giuli, D.: A Context Model for Context-Aware System Design Towards the Ambient Intelligence Vision: Experiences in the eTourism Domain. In: Stephanidis, C., Pieper, M. (eds.) ERCIM Ws UI4ALL 2006. LNCS, vol. 4397, pp. 173–191. Springer, Heidelberg (2007)

3. Han, M., Kim, G., Park, S., Kim, L., Ha, S.: Process and Location-Aware Information Service System for the Disabled and the Elderly. In: Stephanidis, C. (ed.) UAHCI 2009. LNCS, vol. 5616, pp. 207–215. Springer, Heidelberg (2009)
4. Rodríguez, M.D., Favela, J., Martínez, E.A., Muñoz, M.A.: Location-Aware Access to Hospital Information and Services. IEEE Transactions on Information Technology in Biomedicine 8(4), 448–455 (2004)
5. Ongenae, F., Strobbe, M., Hollez, J., Jans, G.D., Turck, F.D., Dhaene, T., Demeester, P.: Ontology Based and Context-Aware Hospital Nurse Call Optimization. Complex, Intelligent and Software Intensive Systems, 985–990 (2008)
6. Patient Guidance System Featuring Electronic Paper Technology, `http://pr.fujitsu.com/jp/news/2009/10/5-3.html`
7. Beyer, H., Holtzblatt, K.: Contextual Design: A Customer-Centered Approach to Systems Designs. Morgan Kaufmann, Los Altos (1997)

Impact of Distance to Screen upon Spacial Awareness

Kazuya Kuroda[1] and Makio Ishihara[2]

[1] Graduate School of Computer Science and Engineering, Fukuoka Institute of Technology, 3-30-1, Wajiro-higashi, Higashiku, Fukuoka, 811-0295, Japan

[2] Department of Computer Science and Engineering, Fukuoka Institute of Technology, 3-30-1, Wajiro-higashi, Higashiku, Fukuoka, 811-0295, Japan

mfm10007@bene.fit.ac.jp , m-ishihara@fit.ac.jp

Abstract. In this paper, we conducted an experiment on what impacts the distance between the user and computer screen as well as the size of FOVs would give on the spacial awareness of 3D virtual worlds. One of the interesting findings is that the distance between the subject and computer screen plays an important role in the spacial awareness of 3D virtual worlds. The spacial awareness is improved when we see the computer screen in the distance.

1 Introduction

In recent years, the increase of computers' processing speed is significant and makes it possible to draw 3D virtual worlds precisely in real time. Second-Life of social networking service or SNS sites, and "Call of Duty" and "Metroid Prime" of first-person shooters on Nintendo Wii or Sony PlayStation are some of good examples. Even mobiles such as iPhone and Android devices can deal with 3D virtual worlds. 3D virtual worlds are now commonplace. In order to draw a 3D virtual world we often use a virtual camera. A virtual camera is an imaginary camera in a virtual world, which is used to generate screen images. Then we observe these images on a computer screen. When it comes to field of views or FOVs, there is a concern about inconsistency between them. That is, there are two FOVs. One is of the virtual camera and the other is of our eyes looking at the computer screen. The FOV of our eyes looking at the computer screen is not always the same with the FOV of the virtual camera. As for the size of the FOV of our eyes, Kobori [1] showed that the efficiency of finding a way to the exit in 2D mazes tends to worsen when we see them through a narrow window. The efficiency is defined as a ratio of the shortest path to the path taken by the subject.

Our study is about what impacts the inconsistency between the FOVs as well as the size of them would give on the spacial awareness of 3D virtual worlds. In this manuscript, we conducted an experiment on the impact of three cases of the FOVs, upon the spacial awareness. The first case evaluates the impact of the wideness or narrowness of FOVs. The second one evaluates the impact of the distance between the subject and computer screen with constant FOVs. The third one evaluates the impact of the difference between the FOV of the virtual camera and the one of our eyes. In our experiment, we used some 3D mazes for 3D virtual worlds.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 270–276, 2011.

© Springer-Verlag Berlin Heidelberg 2011

This paper is organized in the following manner. Sect. 2 explains the experimental method then Sect. 3 shows the results. Finally, Sect. 4 gives the concluding remarks.

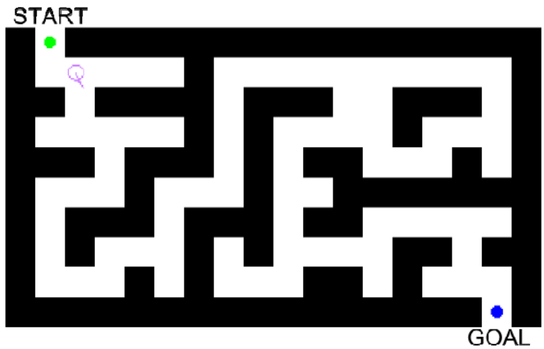


Fig. 1. Top view of one of the 3D mazes used in the experiment

Fig. 2. Screenshot of a subject's view during the experiment

2 Experimental Method

A 3D maze is a system of paths separated by walls in a 3D virtual world, that is designed so that it is difficult to find your way through. Fig. 1 shows one of the 3D mazes used in our experiment. There were 14 subjects between the age of 21 and 22. Each subject was given five mazes from a set of pre-prepared ten mazes at random. Each maze was designed ten blocks wide and 18 blocks long. The ratio of the number of blocks for paths to the number of blocks for walls is set to about 40 percent, making the maze most difficult to be solved [2] [3]. Each subject was asked to solve them on five conditions of the following 15 ones at random. During the experiment, the system tracked the subject's position and orientation.

Table 1. All the 15 conditions of the experiment

Distance \ Our eyes' & camera's FOVs	10°	20°	30°
50cm	1	2	3
100cm	4	5	6
150cm	7	8	9

Camera's FOV \ Our eyes' FOV	10°	20°	30°
10°	-	10	11
20°	12	-	13
30°	14	15	-

- **Condition 1, 2** and **3:** Both the FOVs of the virtual camera and our eyes looking at the computer screen are fixed to 10, 20 and 30 degrees, respectively, and the computer screen is placed 50cm away from the subject.
- **Condition 4, 5** and **6:** Both the FOVs are fixed to 10, 20 and 30 degrees, respectively, and the computer screen is placed 100cm away from the subject.
- **Condition 7, 8** and **9:** Both the FOVs are fixed to 10, 20 and 30 degrees, respectively, and the computer screen is placed 150cm away from the subject.
- **Condition 10** and **11:** The FOV of the virtual camera is fixed to 10 degrees and the one of our eyes looking at the computer screen is fixed to 20 and 30 degrees, respectively, and the computer screen is placed 50cm away from the subject.
- **Condition 12** and **13:** The FOV of the virtual camera is fixed to 20 degrees and the one of our eyes looking at the computer screen is fixed to 10 and 30 degrees, respectively, and the computer screen is placed 50cm away from the subject.
- **Condition 14** and **15:** The FOV of the virtual camera is fixed to 30 degrees and the one of our eyes looking at the computer screen is fixed to 10 and 20 degrees, respectively, and the computer screen is placed 50cm away from the subject.

Table 1 summaries all the 15 conditions. Each condition is numbered 1 to 15. The left part of Table 1 shows conditions where both the FOVs are set to the same. For example, if the FOV of the virtual camera is set to 20 degrees, the one of our eyes is set to 20 degrees as well. While holding this property, it varies in the size of the FOVs and the distance between the subject and computer screen. The right part of Table 1 shows conditions where the FOVs are not set to the same. For example, if the FOV of the virtual camera is set to 20 degrees, the one of our eyes might be 10 or 30 degrees. Along with the conditions, the distance between the subject and computer screen is kept constant. In this experiment, it is 50cm.

3 Experimental Results and Discussion

Fig. 3 shows the search efficiency at each condition across all the subjects. In the figure, the horizontal axes place all the conditions and the numbers correspond to those of Table 1. The z axis shows the search efficiency. The search efficiency is defined as the ratio of the shortest path P_{short} to the path taken by the subject $P_{subject}$:

$$\text{Search efficiency} = P_{short} / P_{subject} * 100 \tag{1}$$

The shortest path P_{short} denotes the number of blocks of the maze, where the shortest path lies. The subject's path $P_{subject}$ denotes the number of blocks where the subject walks. It does not matter if the subject walks on the same block multiple times. The search efficiency tells 100 if the subject takes the shortest path while it tells less if he/she takes longer paths.

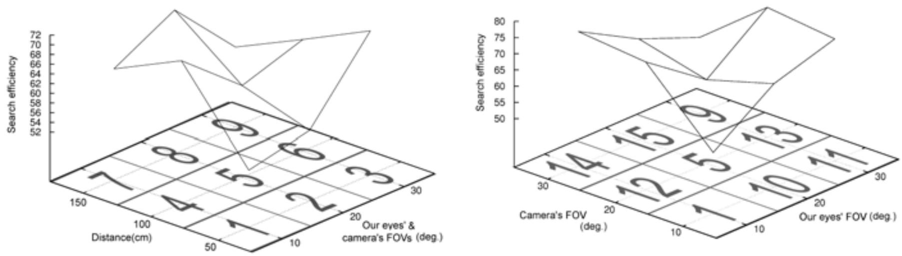


Fig. 3. Search efficiency for each condition

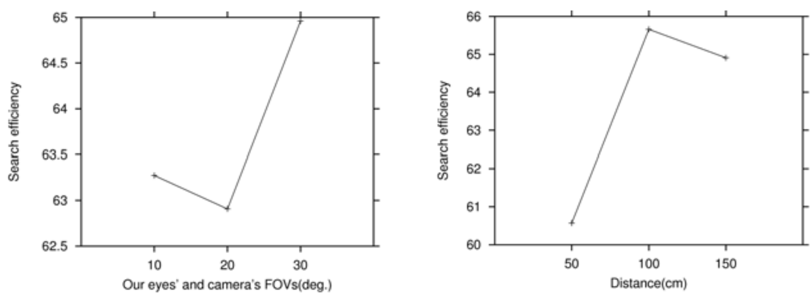


Fig. 4. Average of the search efficiency over the distance(left) and FOVs(right)

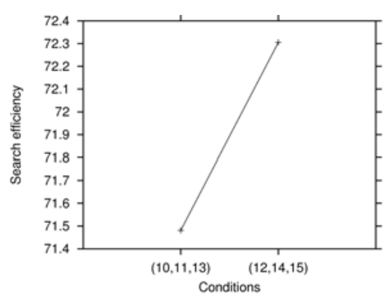


Fig. 5. Average of the search efficiency over the conditions 10, 11 and 13, and the conditions 12, 14 and 15

Fig. 4 is derived from the left graph of Fig. 3 by averaging the search efficiency over the distance and FOVs. The left graph of Fig. 4 shows that the search efficiency increases when the FOVs widen. Moreover, the right one of Fig. 4 shows that the search efficiency increases when the distance between the subject and computer screen becomes long. These results imply that the search efficiency will be improved when we see the computer screen in the distance with wide FOVs. Fig. 5 is derived from the right graph of Fig. 3 by averaging the search efficiency over the conditions 10, 11 and 13, and the conditions 12, 14 and 15. The former features that the FOV of our eyes is wider than the one of the camera while the latter features that the FOV of our eyes is narrower than the one of the camera. From the result, there is a certain difference in search efficiency between these two cases of FOVs. The search efficiency is likely to increase when we see compressed images on a small computer screen.

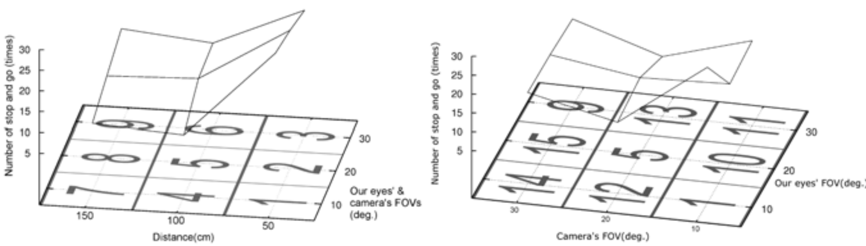


Fig. 6. Effort to solve mazes for each condition

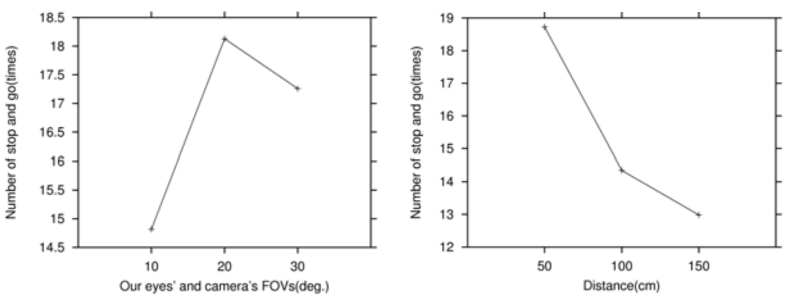


Fig. 7. Average of the effort over the distance(left) and FOVs(right)

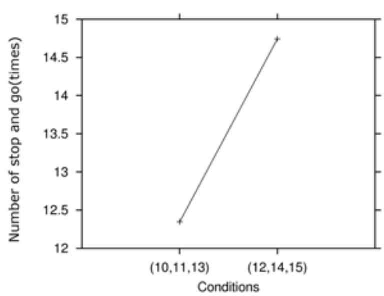


Fig. 8. Average of effort over the conditions 10, 11 and 13, and the conditions 12, 14 and 15

Fig. 6 shows effort to solve mazes at each condition across all the subjects. In the figure, the horizontal axes place all the conditions and the z axis shows the effort. The effort is defined as the number of stop and go during the subject solves each maze. If the number increases, it means that the subject is likely to put effort into recovering the spacial awareness.

Fig. 7 is derived from the left graph of Fig. 6 by averaging the effort over the distance and FOVs. The left graph of Fig. 7 shows that the effort increases when the FOVs widen. The right graph of Fig. 7 shows that the effort decreases when the distance between the subject and computer screen becomes long. These results imply that the subject puts less effort into solving mazes when he/she sees the computer screen in the distance with narrow FOVs. Fig. 8 is derived from the right graph of Fig. 6 by averaging the effort over the conditions 10, 11 and 13, and the conditions 12, 14 and 15. From the result of Fig. 8, there is a certain difference in effort between these two cases of FOVs. The effort is likely to increase when we see compressed images on a small computer screen.

All the findings above are summarized as follows.

1. 1) The search efficiency will be improved when we see the computer screen in the distance with wide FOVs while the subject will put less effort into solving mazes when he/she sees the computer screen in the distance with narrow FOVs.
2. 2) Both the search efficiency and effort are likely to increase when we see compressed images on a small computer screen.

The first sentence says that the distance between the subject and computer screen plays an important role in both the high search efficiency and less effort.

The second one says that we put more effort into solving mazes to improve the search efficiency when we see compressed images on a smaller computer screen. This means that when we design computer games with 3D virtual worlds for mobiles and portable game consoles, we must be aware of a tradeoff between the search efficiency and effort, leading to the settings of virtual cameras. Finally, note that we do not see any relationship between the search efficiency and effort in this experiment. Putting more effort into solving mazes does not necessarily mean making the search efficiency higher.

4 Conclusions

In this paper, we conducted an experiment on what impacts the inconsistency between FOVs as well as the size of them would give on the spacial awareness of 3D virtual worlds. One of the interesting findings is that the distance between the subject and computer screen plays an important role in the spacial awareness of 3D virtual worlds. The spacial awareness is improved when we see the computer screen in the distance. In future work, we are going to study with wider FOVs because the FOV of the virtual camera for common consumer games is set to 60 degrees. It is wider than those in our experiment. We are also going to conduct experiments in real space in order to evaluate the impact of our eyes' depth perception on the spacial awareness because we used a flat computer screen to display 3D mazes in the experiment.

References

1. Kobori, S.: Analysis of Ability of Problem Solving as for Searching in Maze. In: Proc. 42th Annual Convention IPS Japan (2), 269–270
2. Mizusawa, M., Kurihara, M.: Hardness Measures for Maze Problems Paramete-rized by Obstacle Ratio and Performance Analysis of Real-Time Search Algorithms. In: TJSAI, vol. 12(1) (1997)
3. Wang, Y., Matsushita, S.: A study about the space recognition by drawing of a sketch map and scenery after a maze search walk. Summaries of academic lectures of Architectural Institute of Japan E-1, 869–870 (2006)

Why History Matters: Visual Innovation and the Role of Image Theory in HCI

Margarete Pratschke

eikones NCCR Iconic Criticism, ETH Zurich, Science Studies,
Rämistr. 36, 8092 Zürich, Switzerland
pratschke@wiss.gess.ethz.ch

Abstract. Creating and experiencing visual innovation in HCI is based on the historical property of images, and builds on users' visual experiences with certain prior imagery. Historicity of digital visuality seems to be crucial but ignored by most approaches in designing visual experience. This paper suggests to take into account not only aspects from image theory but also art history and science studies that suggest ways of dealing with the historicity of images for successfully creating "new" visual form. Furthermore HCI's own history bears opportunities to connect to the epistemics of visuality derived from the past.

Keywords: History of HCI, Image Theory, Art History, History of Science, Science Studies, Visual Studies, Epistemology, Visual Thinking.

1 How Images Work: A(n) (Art) Historical Approach to HCI

Seeing and understanding certain visual forms, structures, or images intuitively when interacting with technology are key aspects of usability and of successfully creating user experience. The epistemics of visual thinking – perceiving as understanding – is closely connected to image theory and a field of lively debate in various historically grounded disciplines. Although working visually and producing images and iconic visual structures when creating user experiences, HCI is not referring to recent image theory that deals with functions of images in science and technology.

Research within the last two decades from the History of Science, Science Studies, Visual Studies and especially from Art History in the german-speaking countries, has produced deep insights into the ways images work and the role they play in visual strategies in science and technology [3, 16, 4, 8, 7] – especially when the extremely constructive character of any visual form has been emphasized [5]. Images in both art and science have been investigated not only in terms of their aesthetics, but mainly regarding their epistemic functions [6]. As a result the notion of images as purely illustrative elements has been dismissed. In this discourse the understanding of the term 'representation' has been questioned; furthermore it has been concluded that images and visuality play a generative role in the production of knowledge in science and technology [23, 24]. In particular, visual thinking and concepts of implicit knowledge seem to influence both the production of images by artists, designers, and scientists and the reception of images by observers or users [11].

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 277–284, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Another closely connected result of the research on images concerns their historicity, as images always seem to refer to other (preceding) images. The specific momentum and power of images, that may even be contrary to any precise purpose or instrumental use, are historically informed – by the very own image history. Any interpretation, as well as any further "visual innovation" of technical images or visual experiences – as in HCI – have to consider the historicity of the "visual" in general: the form, use and function of images depend on a continuous historical process in which visual traditions and new imaging techniques intersect [7, 15].

2 History of Interacting with Images in Art and Technology

Perceptual traditions shape today's user experience. Such traditions can be traced by comparing recent solutions to historical ones, or by comparing displays to handle visual information made by artists throughout the 20th century to technological solutions [18, 19, 21]. Examples from the modern era in art, i.e. exhibition installations or architectural collages, could serve not only as sources of inspiration for the design of interfaces but rather as fundamental theoretical studies of displaying multiple workspaces interacting with users in a spatial setting.

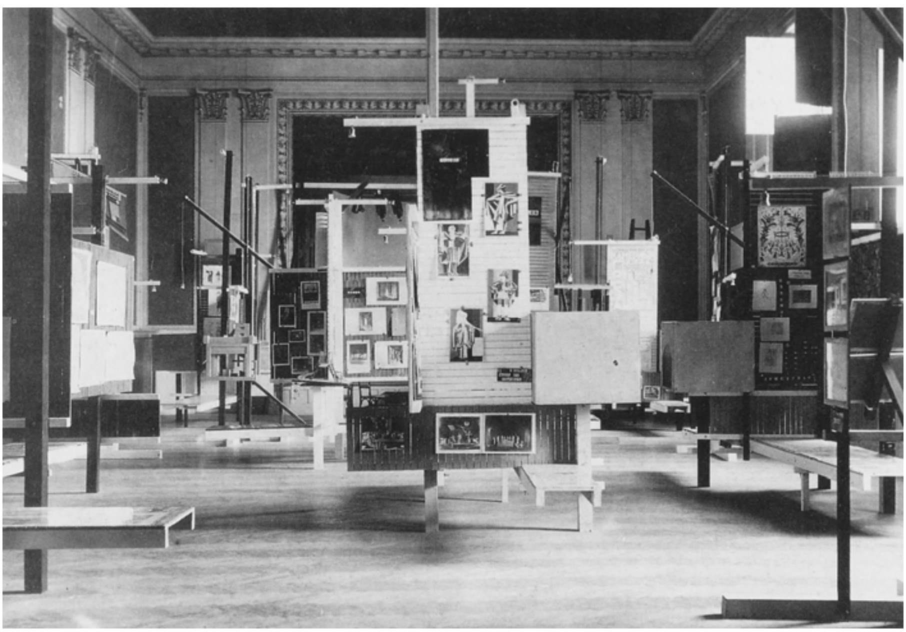


Fig. 1. Frederick Kiesler, Internationale Ausstellung neuer Theatertechnik, Vienna, 1924. Interactive display installation for exhibitions with overlapping structures.

The way in which people interact with multiple, spatially arranged images and objects – in analogy to multiple workspaces on screens or in spatial interaction ('arranging windows') – was vastly theoretically discussed and practically explored in avant-garde architecture and design. The problem of dimensionality, movement and interactivity of viewer/user and objects or images were the main goals tackled by several artists: for instance in the architecture of Mies van der Rohe in the 1930s, or the exhibition design and architecture of Frederick Kiesler in the 1920s (Fig. 1), the films and spatial experiments of Hans Richter and Theo van Doesburg in the 1920s, or Herbert Bayer's display-installations for pictures. But also designers like Ray and Charles Eames who were working in the avant-garde tradition in the 1960s and 70s investigated into the ways of communication and interaction technology – within the art-sphere, but, as in the case of the Eames, were collaborating with and working for corporations like IBM. All of these examples of experimental methods of arranging images in time and space through collage-like interfaces, play with the viewer's sense of disorientation, thus challenging him to interact all the more strongly.

As artistic expressions of user-centered design, such examples could serve not only as sources of inspiration for the design of interfaces but also as fundamental theoretical studies of displaying multiple panels or workspaces interacting with users in a spatial setting. Comparing historical examples of interacting with art or architecture to current ways of interaction experiences in contemporary HCI, does not aim at transforming these images into art [7]; the issue is rather how the interplay of visible forms is intimately connected with the viewers'/users' visual experiences in a diachronic way – and is informed by the history of images in the context of communication, technology, science and art.

"Visuality" has a specific rhetoric, with significant communicative and mnemological laws, not only effective in the sphere of so-called "high art" but in the production, design and interpretation of scientific representations and records as well. Thus images in art and technology likewise build on their own historicity. Based on this historical argument, perceptual traditions and historically imparted visual forms have impact on the design and reception of visual innovation in HCI. By tracing visual conventions resulting from the history of images in both art and technolgy new starting points for visual innovations might be created.

3 Tracing Traditions: Image Theory and Epistemology in HCI

Computer Science and HCI rarely refer actively to their own history when developing future scenarios, though at crucial historical moments of visual innovation in interaction design, the protagonists were not only practically developing new images, but were dealing with image theory derived from art history or were even referring to contemporary contexts of art and design. One aspect of the efficiency of these approaches is characterized by the reference to concepts of visual thinking and the preference of visuality and visibility by their creators.

The historical impact of visual thinking in terms of gestalt theory as well as concepts of implicit pictorial knowledge within computer science, offer another way of tracing the successful creation of conventions that shape today's visual forms in HCI from a theoretical point of view.

3.1 Visual Thinking: Xerox PARC, the GUI, and Rudolf Arnheim

Although several scholars in the history of science have pointed to contacts and productive exchanges of gestalt psychologists Kurt Lewin and Wolfgang Köhler with protagonists of cybernetics [12], as well as to Michael Polanyi's exchange with Alan Turing [25], the influence of Rudolf Arnheim, Ernst Gombrich and gestalt theory on human-computer interaction, artificial intelligence, on the research of Xerox PARC, and on scientists such as Douglas T. Ross, J. C. R. Licklider, Alan Kay, Marvin Minsky, Herbert A. Simon, and Allen Newell has remained unnoticed until recently [20, 21]. What all of them had in common was a belief in the virtue of images and explanations of perception as intuitive cognition.

Rudolf Arnheim had developed a theory of "visual thinking" based on gestalt psychology [2]; in his theory he attacked the assumption that thinking is only possible in words, and that language precedes perception. Quite the contrary he argued that "all perceiving is also thinking, all reasoning is also intuition, all observation is also invention" [1]. In this sense "gestalt" was not conceived of as just a way of defining visual form or shape, but rather – as did experimental gestalt psychologists – as a way of understanding human cognition that is based on visual knowledge. In this notion the concept of gestalt offered an epistemic theory of perception and visuality as a creative process that comprised instantaneous and direct ways of problem solving contrary to cybernetic explanations based on procedural processes in temporal sequences. In fact the theory of gestalt formed a basis for early interaction design, visual interfaces and the shaping of pictorial interface elements like windows, icons, and menues that were developed at Xerox PARC.

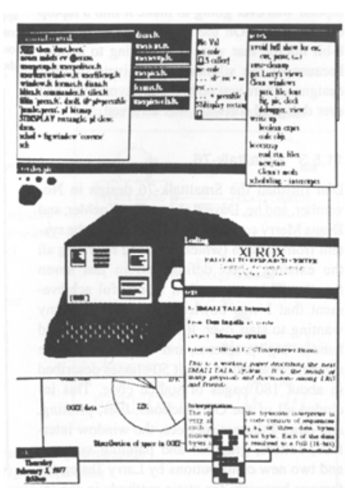


Fig. 2. Smalltalk-72 Interface, Xerox PARC, Alan Kay and Learning Research Group, [14] Fig. 11.51

Xerox PARC's work during the 1970s led to one of the most successful products in the history of HCI, the graphical user interface (fig. 2). Based on the work for the programming language Smalltalk, Alan Kay and the Learning Research Group at

Xerox PARC referred to pedagogical theories as well as art historical image theories by art historian Ernst Gombrich, along with art psychologist and gestalt theorist Rudolf Arnheim [21, 15, 26, 2, 10]. In particular, David C. Smith's thesis [26] demonstrates not only an example of visual programming and of Kay's concept of iconic programming, but an extraordinary discussion and reference to Gombrich and Arnheim in tradition of the gestalt reception in computer science from the 1960s.

Thus Gombrich's theory of stylistic historical evolution of art, that refuted the notion of an "innocent eye", became an arthistorical groundwork of a technological process leading to pictorial human-computer interaction [10]. Furthermore Arnheim's concept of "visual thinking" served as an argument to reject any interaction with computers that was word- or number-based [2].

As a result many elements of the graphical user interface seem to reflect experiments conducted by gestalt theorists and offer historical examples and opportunities for further productive investigation into the trans disciplinary correlations and the diachronic character of visual form (fig. 3, 4).

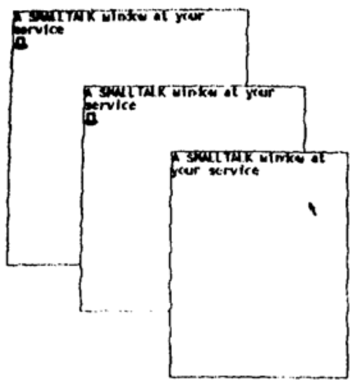


Fig. 3. Early Smalltalk Windows on Interim Dynabook (Alto), Xerox PARC, Alan Kay and Learning Research Group, Smalltalk-72, [14] Fig. 11.32

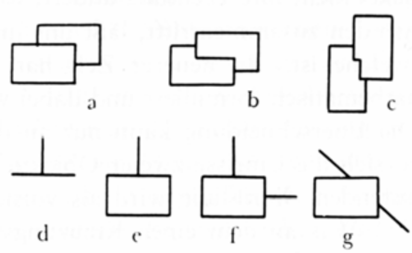


Fig. 4. Gestalt experiments on creation of visual space, 'Depth by Overlapping', Rudolf Arnheim, Art and Visual Perception, [1] fig. 180

3.2 Tacit Knowledge: Douglas Engelbart, the Display, and Herman Miller

While Xerox PARC's experimental work on iconic programming and visual interaction almost lead to a dogmatization of "visibility" when constructing the first commercial graphical user interface for the Xerox Star, a similar outcome resulted from another historical collaboration that crossed the boarders of computer science and art/design. Working on his Augmenting Human Intellect workspace at Stanford Research Institute (SRI) in the 1960s, Douglas Engelbart collaborated with Herman Miller Research, headed by Robert Propst, who was doing research that led to an innovative office system, the "Action Office". Both researchers developed human-centered strategies for "displays" from the respective point of view of art and technology [21, 9, 22].

Engelbart's concept of Augmenting Human Intellect [9] came up with a workstation, interface, and connecting devices for man-machine interaction of the NLS computer system. It was informed by ergonomic considerations resembling those of the designers and artists working at Herman Miller – and bears resemblances to Michael Polanyi's theory of bodily implicit knowledge [17]. Although not explicitly referring to Polanyi, Engelbart's contemporaneous concept of bodily connections of man and machine via mouse, chorded keys or other devices, seem to be concrete deductions of Polanyi's theoretical concept of tacit or implicit knowledge.

Fig. 5. Robert Propst, The Office. A Facility Based On Change, 1968, Herman Miller Research, Action Office System, [22] p. 51

Prior to Engelbart's 1968 live demonstration of the NLS computer system his team collaborated with Herman Miller Research and was even testing a mock up of Herman Miller's Action Office furniture set up at the Augmenting Research Center at SRI. The Action Office was designed by George Nelson and had been the result of experimental efforts by Herman Miller's Research Lab guided by Robert Propst (fig. 5).

Robert Propst had devised a new understanding of the office as a flexible display system, based on "visual image triggers", experimentally tested for the „human performer" who had a „spectacular ...tool: the human eye as a receptor for the mind." [22]. With its concepts of "multi-work stations" and vertically oriented displays one can compare Propst's efforts not only to exhibition installations from the avant-garde in art but also to the preference of visuality in the history of computing that led to early versions of graphical interaction between humans and computers, or: resulted in interactive "displays".

4 History Matters: The Challenge of Intersecting Past and Future

In the light of image theoretical arguments stressing that visuality always refers to its own historicity it seems quite clear that history matters when designing visual innovation. Certain visual conventions, image traditions and the uses of images build on earlier expieriences of any user or viewer. It may be concluded that the historicity of images – in art and design history as well as in HCI's own history – lays the groundwork for further understanding of how images work in designing interaction. Any future visual innovation is based on history.

References

1. Arnheim, R.: Art and Visual Perception. In: A Psychology of the Creative Eye, Univ. of California Press, Berkeley (1954)
2. Arnheim, R.: Visual Thinking. Univ. of California Press, Berkeley (1969)
3. Boehm, G. (ed.): Was ist ein Bild? Fink Verlag, München (1994)
4. Bredekamp, H.: A Neglected Tradition? Art History as Bildwissenschaft. In: Critical Inquiry, vol. 29, pp. 418–428. Chicago Univ. Press, Chicago (2003)
5. Bredekamp, H.: Theorie des Bildakts. Über das Lebensrecht des Bilder. Suhrkamp, Berlin (2010)
6. Bredekamp, H., Schneider, B., Dünkel, V. (eds.): Das Technische Bild. Kompendium zu einer Stilgeschichte technischer Bilder. Akademie Verlag, Berlin (2008)
7. Bruhn, M., Dünkel, V.: The Image as Cultural Technology. In: Elkins, J. (ed.) Visual Literacy, pp. 165–178. Routledge, New York (2008)
8. Elkins, J.: The Domain of Images. In: On the Historical Study of Visual Artifacts, Cornell Univ. Press, Ithaca (1999)
9. Engelbart, D.C.: Augmenting Human Intellect. A Conceptual Framework. SRI Report, Menlo Park, CA (1962)
10. Gombrich, E.H.: Art and Illusion. In: A Study in the Psychology of Pictorial Representation. Pantheon Books, New York (1960)

11. Hagner, M.: Sehen, Gestalt und Erkenntnis im Zeitalter der Extreme. Zur historischen Epistemologie von Ludwik Fleck und Michael Polanyi. In: Bader, L., Gaier, M., Wolf, F. (eds.) Vergleichendes Sehen, pp. 337–357. Fink Verlag, München (2010)
12. Heims, S.J.: Constructing a Social Science for Postwar America. In: The Cybernetics Group, pp. 1946–1953. MIT Press, Cambridge (1993)
13. Kay, A.: User Interface. A Personal View. In: Brenda, L. (ed.) The Art of Human-Computer Interface Design, pp. 191–207. Addison-Wesley, Reading (1990)
14. Kay, A.: The Early History of Smalltalk. XI: Smalltalk Session. In: Bergin Jr., T.J. (ed.) History of Programming Languages II, pp. 511–598. ACM, New York (1996)
15. Kemp, M.: Visualizations. The Nature Book of Art and Science. Oxford Univ. Press, Oxford (2000)
16. Maar, C., Burda, H. (eds.): Iconic Turn. Die neue Macht der Bilder, DuMont, Köln (2004)
17. Polanyi, M.: The Tacit Dimension. Routledge & Kegan Paul, London (1966)
18. Pratschke, M.: Jockeying Windows. The pictorial space generated by graphical user interfaces as the visual basis for multitasking. In: Neue Gesellschaft für Bildende Kunst, Multitasking. Synchronität als kulturelle Praxis, pp. 7–8. NGBK, Berlin (2007)
19. Pratschke, M.: Interaktion mit Bildern. Digitale Bildgeschichte am Beispiel grafischer Benutzeroberflächen. In: Bredekamp, H., Schneider, B., Dünkel, V. (eds.) Das Technische Bild. Kompendium zu einer Stilgeschichte technischer Bilder, pp. 68–81. Akademie Verlag, Berlin (2008)
20. Pratschke, M.: Die grafische Benutzeroberfläche als Bild. Zur Rezeption von Rudolf Arnheim und Ernst Gombrich in der Computer Science der 1970er Jahre. In: kritische berichte. Zeitschrift für Kunst und Kulturwissenschaften, vol. 37/4, pp. 54–63. Jonas Verlag, Marburg (2009)
21. Pratschke, M.: Windows als Tableau. Zur Bildgeschichte grafischer Benutzeroberflächen. PhD thesis, Humboldt-Universität zu Berlin (2010)
22. Propst, R.: The Office. In: A Facility Based on Change, Herman Miller, Zeeland, Mich (1968)
23. Rheinberger, H.-J.: Toward a history of epistemic things. In: Synthesizing proteins in the test tube, Stanford Univ. Press, Stanford (1997)
24. Rheinberger, H.-J.: Experimental systems - graphematic spaces. In: Lenoir, T., Beer, G. (eds.) Inscribing Science. Scientific Texts and the Materiality of Communication, pp. 285–303. Stanford Univ. Press, Stanford (1998)
25. Schaffer, S.: OK Computer. In: Hagner, M. (ed.) Ecce Cortex. Beiträge zur Geschichte des modernen Gehirns, pp. 254–285. Wallstein, Göttingen (1999)
26. Smith, D.C.: Pygmalion: A Computer Program to Model and Stimulate Creative Thought. Birkhäuser, Basel (1977)

Image, Imagination, Innovation: How Can We Overcome the Common Ground in the Processes of Visual Design?

Michael Renner[1,2]

[1] Visual Communication Institute, The Basel School of Design, HGK FHNW
[2] eikones, National Center of Competence in Iconic Research, Basel, Switzerland
michael.renner@fhnw.ch

Abstract. This paper presents the analysis of two poles in the spectrum of image creation: the primal drawing process and the coding process in the field of generative design. The reflection upon these design processes is conducted in order to answer the question of how it is possible to overcome individually, socially and culturally determined visual schemata.

Keywords: Innovation, Graphic Design, Design Process, Visual Communication, Iconic Research, Practice Led Iconic Research, Generative Design, Interface Design, User Experience, Cognition, Anthropology.

1 Introduction

Mankind is described as a cultural being, which specifies the pronounced human ability to develop new achievements relevant for a society and its' prosperous further development over generations. The phylogenetic development of language has often been referred to as proof of the unique capabilities of the human being to develop beyond genetically determined capabilities. Therefore it is not surprising, that we find a wide range of recent publications presenting research results about the origin and the development of human language. [1] [2]

In contrast to the anthropological and linguistic studies, we find little reflection about innovation processes involved in the creation of images. If we refer to their history we find a few indications. Plinius the elder, a Roman writer from the first century AD, describes for example how the painter Zeuxis was commissioned by the people of Agrigentum to create a picture of the goddess Juno and therefore chose the most beautiful women in the town to combine their most beautiful features in his painting [3]. This report is proof that already in the first centuries AD the artist was described as a responsible individual able to determine analytically an aesthetic quality.

In the Renaissance, Federico Zuccari describes the artist as an individual following the divine processes of creation by imitating nature and moves the process of imagination and creation beyond human abilities. [4]

Summarized under the term creativity, some references with the aim to induce creative behavior in Management, Art, Advertising or the scientific context can be found

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 285–294, 2011.
© Springer-Verlag Berlin Heidelberg 2011

in the middle of the 20th Century[1]. Most of these authors summarize their results in simplified formulas applicable to practical situations.

More recent publications in the Theory of Sciences show that there is not a simple pattern which we can apply to a field of scientific innovation. [5] The study of modern art history confirms that we can only look back at innovative developments after they were achieved.

It is the focus of the research presented here to discover the implicit (or tacit) knowledge involved in the process of generating unseen visual constellations for our daily communication. This research is situated in the context of *Iconic Research*[2] with the specific purpose of gaining knowledge for those who create images and are confronted with the lack of scientific reflection about the potential of images to generate meaning. The key question stays the same in the context of language, of science, of technology or visual communication, design and art. Mark Johnson describes the general direction of this inquiry as follows:

“Our ability to make new meaning, to enlarge our concepts, and to arrive at new ways of making sense of things must be explained without reference to miracles, irrational leaps of thought, or blind impulse. We have to explain how our experience can grow and how the new can emerge from the old, yet without merely replicating what has gone before. As it turns out, this may be one of the most difficult problems in all of philosophy, psychology, and science: how is novelty possible?” [6]

If we try to define the question more precisely for the context of visual communication, we can refer again to the anthropological studies concerned with the development of human communication. They describe the *common ground* as a pre-requisite for the most basic occurrences of gestural communication. [7] Even a simple *pointing gesture,* executed by a person to guide the attention of another individual to an object or event in the context of communication, can only generate meaning if both individuals know from each other what they know. As described by Tomasello, if I am in a room and point at the door, addressing another person, it can mean different things. If I have spent hours in the room in a meeting, the gesture could mean: “Lets go outside, I have been in here long enough.” If we just arrived it could mean: “Lets close the door and start the meeting.” Depending on the situation the pointing could also demand from the addressed person to leave the room.

For sharpening our questions of innovation in the context of images we can derive from the anthropological studies a similar necessity of a *common ground* to understand images. We need to know what a house is and how it looks like in order to be able to recognize it in an image. How is it possible, under these circumstances, to arrive at an image of a house which goes beyond the five lines of the pictogram culturally embossed in our mind? What are the processes to generate a new visual form of a house, which is recognized, accepted and understood in our daily communication with images? In order to analyze the processes of image generation it makes sense to start with the basic methods of image generation – the drawing process.

---

[1] For example: Alex F. Osborne, Creative Education Foundation, University of Buffalo.

[2] Iconic Research is a research field, which developed out of the observation of the *Pictorial Turn* (Mitchell 1994) and *Iconic Turn* (Boehm 1995).

2 Drawing as an Exemplary Methodology to Create Unseen Images

2.1 Cognitive Linguistics as a Foundation for the Analysis of Innovation Processes in the Context of Drawing

To draw, is a physical process where the movement of the hand holding a tool leaves a mark on a surface. If we try to describe the process of drawing precisely, we discover that there are parts of the process which we can address easily with words. For example we can describe the conscious definition of criteria which makes it possible to judge the final drawing in the preparation phase. We can say why we choose charcoal instead of a pencil. If we go more in detail, we will discover, that the decisions made in the main motor activity of the process, the drawing of the line, are not describable through our language. In other words the experiment of describing the drawing process in detail shows the interaction between decisions made consciously and those executed under the level of consciousness. [8]

The theories of cognitive linguistics, especially the theories of the embodiment, are describing a continuum between unconscious reactions of an organism and the ability to think in abstract symbolic systems such as language or math. Recent studies have drawn attention to the relevance of the fact that all the human abilities to think in abstract concepts such as language or mathematics have a direct relationship to the physical constitution of the human being and to our mobility in space. The constant reaction of our organism to new events in the environment creates a flow of unconscious emotional activity in our body. These reactions are the foundation of feelings, which we can define as distinguishable qualities of emotions. Feelings are also participating in the creation of image schemas, which are described as multi-modal records of experiences including feelings and sensory aspects of an event as well as basic structural qualities. The ability to recognize the qualities of emotions, observe their patterns, and develop image schemas, is leading to a next level of abstraction through basic metaphors. These metaphors connect two different types of experiences in the form of image schematic objects, perceived through the human body. For example, the experience of passing time is directly related to the human body moving through space. The past is perceived in western cultures as objects we have in back of us. The future is perceived as objects in front of us. The present is where we currently are on the path. A closer look at language shows already in this example that our expressions are reflecting these basic metaphoric ideas and prove their importance in the way we are forming complex concepts. In the context of the basic metaphor we formulate in English: I look forward to see you, he left his past behind him, let's move on, etc… What has been described here as a basic theory, relevant for linguistics, raises the question about the status of images in the continuum from emotional thinking (unconscious cognition) to thinking in abstract symbols. [9]

This simplified summary allows us to analyze the drawing processes based on basic principles outlined above. The decisions made under the level of consciousness, which are not accessible through words, can be addressed in experimental settings and we can try to make the influences on these decisions visible with practical experiments.

2.2 Drawing Experiments I; Drawing an Ellipse

If we look at the drawing process with the aim to understand influences on the unconscious decisions made continuously in the process of this specific method of image generation, we can refer to one of the known academic drawing approaches developed in the French academies in the 19th century. The large format analytical drawings are based on the understanding of the two-dimensional representation of three-dimensional geometric objects such as the cube, the cylinder, the conus or the sphere. Investigating the way these shapes are generated we can pick one basic form and analyze its possibilities to be generated. The form of the ellipse can be generated geometrically by calculating every point of the line or it can be drawn today as one definite line in a digital drawing program. In this context unconscious decisions are not intended and we avoid them as much as possible.

Already the tracing of an existing ellipse by hand through a semi-transparent paper, allows us to examine the behavior of the sensorimotor apparatus. The aim of the movement is to keep the tip of the pen as close as possible on top of the existing line. The loop of movement, setting the mark, perception of the mark, and feedback resulting in the adaption to the movement, leads to the line with a specific quality. We would describe the quality of a traced line as shaky, insecure or tentative.

If we draw an ellipse free hand by developing a definite form through the accumulation of many tentative lines, we are employing the motor abilities of the arm. The resulting line can be described as continuous, showing an aesthetic appeal through its tension. In this very specific drawing technique of an ellipse emerging "out of the paper" through the layering of many lines, we can understand the interdependence of form with the abilities of the sensorimotor system of the human body. We can conclude from this simple observation about the ellipse drawing, that the physical constitution of our arms has a significant influence on form in the drawing process.

2.3 Drawing Experiment II; Mirrored Lines

The following experiment shows a more intricate characteristic of our sensorimotor apparatus. If we create a freely drawn line with the writing hand and try afterwards to draw the same line mirrored with the other hand, we will experience the difficulty to recreate a line with a similar quality as the original.

The attempt to create the mirrored constellation of two lines by drawing with both hands simultaneously leads almost automatically to an unexpected resemblance, when we concentrate on the movement of the writing hand. This example shows, that our sensorimotor system is preferring mirrored movements as long as they are executed simultaneously. This constitution has also an influence on the development of form on the level of unconscious decisions. [10]

2.4 Drawing Experiment III; The Influence of Practice on the Unconscious Decisions

Unconscious decisions become very apparent when we try to create a drawing within a short time span. If we have no chance to reflect about the qualities of the drawn lines, the decisions are staying in the unconscious realm.

The following series of drawings (Fig. 1.), created under different time restrictions, show how practice influences the unconscious decisions in the drawing process. The first drawing of a snapdragon was created within ten seconds, with the aim to represent the flower realistically and shows a clear orientation on the silhouette line. There are hardly any indications of the internal structure of the object and we would not recognize the specific flower in the single drawing. The following drawings created in a longer time span of one minute, ten minutes and a half an hour show, as expected, more structural detail and a clearer representation of the three dimensional special situation. The drawing created within a half an hour shows textural information and indications of how light and shadow are supporting the representation of space. The fifth drawing in the series was again conducted under the timely restrictions of 10 seconds. Like in parallel drawing experiments with the described set up, conducted with other individuals, we can clearly see in the fifth drawing of the presented series the influence of the previous drawing experiences reflected in the unconscious decisions. These experiments are evidence for the strong influence of practice and experience on the unconscious decisions made in the drawing process.

Fig. 1. Indre Grumbinaite, drawing series consisting of 5 drawings of the Snapdragon-flower. Left: First drawing created within 10 seconds. Right: Fifth drawing created within 10 seconds. Middle from top to bottom: Second drawing conducted within one minute; Third drawing conducted within 10 minutes and the fourth drawing conducted within 30 minutes. Format of each drawing: A3, 42 x 29.7 cm, Archive the Basel School of Design HGK FHNW.

2.5 Interpretation of the Drawing Experiments

The described experiments in the area of drawing have isolated some influences on the unconscious decisions made in the drawing process such as the physical ability to move a pencil over a sheet of paper, characteristics of the sensorimotor apparatus and the influence of practice and experience. It is obvious that there are many more factors we can address in this kind of inquiry.

Nevertheless the basic exploration of the drawing process lets us deduce the hypothesis that through the interaction between analytical reflection and emotional reaction, unexpected and unseen images emerge. In this process of interaction, we are able to extend a visual field of options beyond the common ground and beyond the approachable variation accessible through rational extrapolation. Furthermore the examples show that the idea of the accidental in the sense of randomness, is not describing what the drawing process entails. The pre-verbal perception, in the realm of the unconscious, is influenced by the described factors of the human body, its sensorimotor apparatus, the body of society and the cultural context. The processes beyond consciousness are partially addressable in retrospect but can not be anticipated with a symbolic system such as language or math.

3 Generative Design and the Attempt to Create Unseen Images

3.1 Generating Images through Parameters Set by Complex Symbolic Systems

If we look at the opposite spectrum of image generation and turn away from the primal, nonverbal process of drawing to the generation of images through programming code, we can find an early example of images defined by a set of parameters in Laszlo Moholy-Nagy's telephone drawings. He wrote about them:
"Painting by hand may maintain it's historical relevance; sooner or later it will loose it's exclusiveness. […] 1922 I ordered in a street sign factory five pictures out of porcelain-enamel by phone. I had the color samples of the factory in front of me and sketched the images on graph paper. On the other end of the line the head of the department had the same graph paper in front of him which was divided into squares. One of the images was delivered in three different sizes, therefore I was able to study the fine differences in the color relations which were emerging through enlargement and reduction." [11]

We can argue that the innovative image has been generated again in the sketching process on the graph paper and has been "objectified" through a description in coordinate positions transmitted verbally on the phone. We find a similar procedure in Sol Lewitts wall drawing processes which he started in 1968. They consist of a simple set of verbally defined instructions and their execution through a draftsman for a specific situation[3]. The openness of the instruction leads again to an interplay between a seemingly precise verbalized description of the outcome of the wall drawing and the concrete execution of the instructions.

In his statement "Sentences on Conceptual Art" (1969) Sol Lewitt contradicts the description of his method of creating images as a predictable and rationally accessible process of innovation in the first three of his 35 statements.

[3] Sol Lewitt; Wall Drawing #69; 1971; „Lines not long, not strait, not touching, drawn at random using four colors, uniformly dispersed with maximum density, covering the entire surface of the wall."

"1. Conceptual Artists are mystics rather than rationalists. They leap to conclusions that logic cannot reach.
2. Rational judgements repeat rational judgements.
3. Irrational judgements lead to new experience." [12]

In these sentences Sol Lewitt declares his way to escape common ground. The leap to conclusions that logic cannot reach is achieved by irrational thought influenced by emotional judgment in both parts of the process, the formulation of the parameters with language, and the interpretation of the instructions by the person executing the drawing.

Since the first images produced through a programming language, we find the debate about their status. Max Bense called programmed images "artificial art" when he saw the exhibition of Georg Nees at the Technical University in Stuttgart showing programmed images in 1965 [13].

"Overall we can maybe formulate, that the 'artificial' production category distinguishes itself from the 'natural' one through the introduction of an intermediation schema between designer (creator) and image (opus) consisting of program and programming code. With this closely connected is an unusual separation of the work in the aesthetic process." [14]

Like Sol Lewitt has separated the instructions and the execution of the single variation of the wall drawing, generative design (Generative Computergrafik) splits the creative process in the creation of the program and its execution through the computer. Even though both processes are showing this comparable separation, the execution through the computer can hardly be seen as equivalent to the manual execution of an image as described in the wall drawings of Sol Lewitt.

3.2 Analyzing a Generative Design Process

The described split of the creative process in setting the parameters and executing them can be differentiated further if we analyze a concrete example. I will refer to a generative design experiment from 1997, conducted at the Basel School of Design HGK FHNW using Macromedia Director and Lingo[4] [15]. In the first part of the creative process the designer chose the programming environment. This already entails advantages and disadvantages for realizing specific procedures to generate the visual outcome. Usually the languages manageable by visually oriented people are on the level of scripting languages. This lets us deduce that the influences on the design process are divided up even more than Max Bense has declared. The Designer coding his image in a scripting language is already setting his options within a given framework of the more abstract programming language.

In the concrete example a program has been written to generate a sequence of forty circles rotating around the center of a square background (Fig. 2.). The writing of the code is obviously not comparable to the drawing process and its interaction between conscious and unconscious decisions. When we are programming, we are forced to use a highly abstracted vocabulary of the specific language describing procedures

[4] I am fully aware that since 1997 the possibilities of creating generative design have evolved and improved. The principles have not changed since the George Nees has coined the term "Generative Computergrafik" in the 1960's.

based on precise mathematical operations the computer is able to execute. In this situation the programming language is more restricted to a convention than in our natural language where new combinations of words can generate new meaning. New visual constellations appear in the programming context through the programming structure itself as well as through the alterations of variables in the program structure defining the visual outcome. We could argue that the program structure is on a level of abstraction before the actual script has to be written, allowing us to generate unseen constellations, by altering the structural parts. But we do not see any visual outcome just conceptualizing a program structure. The achievement of unexpected visual constellations is often hindered in the phase of realization when the absolute need of mathematical precision occurs. [16]

After the script is running, we can focus in a next phase of the process on the execution under varying conditions by changing the variables influencing the visual outcome. In order to produce new visual constellations on this level, we can generate arrays of numbers with diverse criteria. In the example presented here, the script was set up as an interactive application in which the user generates the variables through the movement of the mouse. Position and distance of the mouse to the center of the circular movement, as well as acceleration and deceleration of the mouse movement, have been used to generate variables. These are employed in the execution of the script to generate variations within the framework of the circular movement script. On this third level of the creative process, we operate at least within the third set of frameworks generated beforehand through other people.

It is remarkable in the presented experiment what meaning the translation of the physical mouse movement and its visual interpretation causes in the beholder. The natural behavior of the elements in this simple experiment and its imitation of a living organism became the springboard for further experiments in 1998[5].

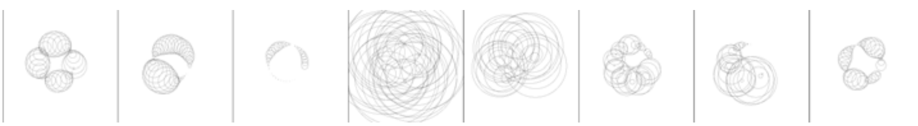


Fig. 2. Roland John: One example from a series of generative design sketches which led to the idea of the “Living Interface”; 1997, Basel School of Design HGK FHNW

3.3 Applying the Generation of Variables through the Mouse Movement to Interface Design

The procedural findings described above in the program structure and the involvement of the user input to generate variations on movement and appearance of the visual elements, were extended in a series of studies towards the design of menu structures for an interface.

[5] For further reference on the theme of cognition and interactive images: Renner M.: Some Thoughts on the Interactive Image, in: Neshan, Iranian Design Magazine Number 22, 2010, Spring/Summer, E: 18 - 22 S. Farsi: 49 - 55 S.

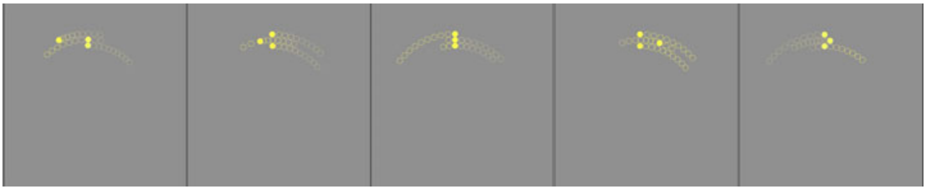


Fig. 3. Roland John: Preliminary interface study for “The Living Interface”; 1998, The Basel School of Design HGK FHNW

From these preliminary experiments with moving interface elements, an interface was designed which allows the access to complex data (Fig. 4.).

Fig. 4. Roland John: Interface Study “The Living Interface”; 1998, The Basel School of Design HGK FHNW

The study shows an exemplary process. In the first investigation on generating visual constellations the analysis led to a set of variations presenting a strong impact on the user, through the natural reaction of the graphic elements on the actions of the user. This analysis has been transposed and overlapped with the informational requirements of an interface. We can find in the combination of two diverse seemingly incompatible schemes, another method of generating innovative functional and visual constellations.

4 Preliminary Conclusion

The comparison of the drawing process with the process of generative design can be analyzed from the perspective of cognitive linguistics as summarized under 2.1.

The unconscious processes of drawing are handled to a large extent on a nonverbal, emotional and sensorimotor level and generate innovative constellations through its conscious employment before the analytical thinking in complex symbolic systems takes over. In opposition, generative design has to employ the complex symbolic system from the beginning in the form of a programming language. These are removed, but derived from the primary sensory experience and are used to generate once again a visual stimuli. This detour from the sensory perception to the metaphoric system of language and back to a visually perceivable stimulus generates visual innovation. The innovative step happens through the abstraction in the verbally described program structures and their openness to derive from them a concrete image by the draftsman (Sol Lewitt) or the programming designer (generative design). [17] Another principal of innovation can be found in the combination of unexpected metaphorical concepts. The overlapping of the two concepts natural movement and informational requirements of an interface can be used quasi as collage to generate new visual meaning.

To refer back to the beginning of this paper and it's title, we can conclude that we can leave preconceived conventions – the common ground – by making use of the blurred border between unconscious and conscious thinking or through the diffusion of translating descriptions of a complex symbolic system to the level of a visually perceivable stimulus.

Because research about the design process has just begun, it is too early to draw any conclusions in this field of inquiry. Nevertheless we can expect that the practice-oriented direction of design research in the form of a Practice Led Iconic Research will contribute to generate a better foundation for the practical field of visual communication. This will allow us to eventually call this field of practice a discipline with its own knowledge base.

References

1. Tomasello, M.: The Origins of Human Communication, Cambridge (2008)
2. McNeill, D.: Gesture and Thought, Chicago (2006)
3. Pliny the Elder: Natural History. Book XXXV, Chicago 64 (2004)
4. Zuccari, F., Kemp, W.: Disegno. Beiträge zur Geschichte des Begriffs zwischen 1547 und 1607. Marburger Jahrbuch für Kunstwissenschaft 19, 232 (1974)
5. Rheinberger, H.-J.: Experimentalsysteme und epistemische Dinge. Eine Geschichte der Proteinsynthese im Reagenzglas, Göttingen, 145 (2002)
6. Johnson, M.: The Meaning of the Body, Chicago, p. 13 (2007)
7. Tomasello, M.: The Origins of Human Communication, Cambridge, p. 78 (2008)
8. Renner, M.: Die stumme Kritik des Entwurfs. Rheinsprung, 11(1) (February 24, 2011), `http://www.rheinsprung11.ch`
9. Lakoff, G., Johnson, M.: Philosophy in the Flesh. New York (1999)
10. Renner, M.: Die stumme Kritik des Entwurfs. Rheinsprung 11(1), 112–114 (2011), `http://www.rheinsprung11.ch`
11. Wick, R.: Bauhaus Pädagogik, Köln, p. 115 (1982)
12. Lewitt, S.: Sentences on Conceptual Art. In: 0-9, New York (1969)
13. von Herrmann, H.-C.: Künstliche Kunst – eine strukturalistische Tätigkeit. In: Nees, G. (ed.) Generative Computergrafik, Berlin (2006)
14. Bense, M.: Aesthetica. Einführung in die neue Aesthetik, 2nd edn., p. 337. Baden-Baden (1982)
15. Renner, M.: Visual Communication with Interactive Media. Swiss Typographic Magazine 6, 22–23 (1999)
16. Reas, C.: Software Structures (2004), `http://artport.whitney.org/commissions/softwarestructures/text.html` (February 24, 2011)
17. Boehm, G.: Die ikonische Differenz. Rheinsprung 11(1), 170–176 (2011), `http://www.rheinsprung11.ch`
18. Boehm, G.: Die ikonische Differenz. Rheinsprung 11(1), 170-176 (February 24, 2011), `http://www.rheinsprung11.ch`

An Extensible Tool for the Annotation of Videos Using Segmentation and Tracking

Marc Ritter and Maximilian Eibl

Chemnitz University of Technology
Chair Media Computer Science
09107 Chemnitz, Germany
{marc.ritter,eibl}@informatik.tu-chemnitz.de

Abstract. Due to massive amount of data, the description of audiovisual media by metadata nowadays can benefit by the support of (semi-)automatic methods during the annotation process. The presented tool enables the user to mark, interactively segment and track preselected objects. An integrated shot detection splits the video into disjoint parts, for instance to circumvent the semi-automated tracking of objects across shot boundaries. Arbitrary application dependent custom image processing chains can be created in conjunction with the research framework AMOPA. Created data is exported in compliance to MPEG7-DAVP.

Keywords: Annotation tool, Image and video processing, Workflow, Object segmentation and tracking.

1 Introduction

In the past years, the application of audiovisual media has grown rapidly among all types of media archives, multimedia centers, knowledge management and e-learning systems, producing a vast and steadily increasing amount of data. Making this information overload searchable becomes more and more challenging concerning time and capacity. The main goal of our works is to decrease the user's workload coming with the annotation process. This goal is successively being achieved by automating the single steps of the workflow.

Machine learning and statistical pattern recognition is engaged in the fields of audio, video and image processing in order to develop classifiers capable of detecting (arbitrary) objects. Unfortunately, the inherent and inevitable training process of a classifier usually demands a huge set of annotated training data, containing examples of the targeted objects. These intellectual annotations are time-consuming and repetitive, and they require a lot of human interaction and attention. Easy, fast and reliable annotation processes may speed up the development cycles of algorithms as well as strengthen a hypothesis of scientific methods when applied on larger data sets.

1.1 Related Work

During the last decades, a lot of different tools have been developed to ease the annotation of images and videos. Successfully applied on the *TREC Video Retrieval*

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 295–304, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Evaluation campaign [1] the ViPER toolkit [2] enables the user to mark single objects in videos, whilst creating descriptors by assistance of a schema editor. In *Caliph* the context of an image can be depicted using a directed semantic graph. Its nodes describe objects and the edges describe the relation between them. Unfortunately, the position of objects cannot be denoted. Its co-partner *Emir* retrieves similar images by means of MPEG-7 low-level descriptors [3]. *Vezzani et al.* train and detect objects on videos while also relying on low-level descriptors [4]. Descriptors of user-selected regions within images are extracted within the *M-OntoMat-Annotizer* to determine objects and their properties.

To minimize the error rate of demanding intellectual annotation the *Multimedia Analysis and Retrieval System* (*MARVEL)* substitutes the graphical user annotation interface against methods from the area of statistical analysis. These methods sort images automatically into categories of a complex taxonomy [6]. A similar approach is proposed by the *VideoAnnEx Annotation Tool* by providing predefined dictionaries for the annotation of key objects, events and static scenes [7].

Structured knowledge can be modeled by extensible XML dictionaries containing categories that can be used for tagging frames or shots [8]. To facilitate frame based annotation, algorithms for semi-automated segmentation and tracking of objects by active contours have proven valuable [9]. *Goldman et al.* initially annotate the shape of objects to interactively track and manipulate its shape within a video [10].

Originally developed to annotate spoken dialogues, *ANVIL* [11] now supports the annotation of diverse video coding schemes. It also integrates XML based export of generated metadata for post processing by toolboxes like SPSS. *Kipp* exploits these coding schemes by adding a spatiotemporal function to mark objects beyond several frames. Multiple sources of video and various sensor data can be processed by the Mac tool *VCode and VData* [12] allowing marking of objects at a timeline. Finally, the consistency of annotations might be checked using integrated verification functions.

A different approach is followed by the tool *FilmEd* [13], which grants multiple annotations from various people simultaneously when connected via a heterogeneous network. Even mobile devices start to be equipped for the annotation process but with very obsolete functions [14].

1.2 Motivation

A lot of efforts have been done to facilitate the cost-intense process of *frame by frame* annotation. On the one hand both object segmentation and tracking could be automated. Unfortunately this option is rarely supported. On the other hand multiple images of a sequence might be tagged with a common semantic concept or item. Albeit the tag itself may be restricted due to the accuracy and correctness of the results of machine preprocessing.

Fan et al. try to close the semantic gap between low-level features and high-level concepts by identifying salient objects within images and matching them with respect to base-level atomic image concepts [16]. A simplified approach is worth striving for

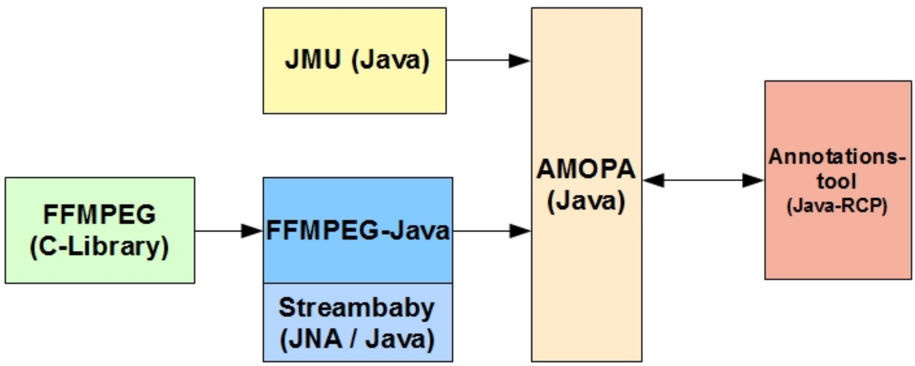


Fig. 1. Architecture of the annotation tool (right). Data is exchanged with AMOPA (middle), that aggregates the functionality of JMU and Streambaby. (from [15])

in any video annotation process. Furthermore, tools potentially lack of the choice of data export to (post) process data individually, depending on the objective of the application.

The here presented annotation tool combines the previously mentioned advantages, while providing the export of annotations to MPEG7-DAVP and the opportunity to model and create flexible and extensible application-dependent processing chains by interconnecting the annotation tool to the research framework AMOPA [17]. Thus, a possibility to integrate fast multi-threaded processing chains is supplied beside methods for rapid structural decomposition of videos as well as marking, segmentation and tracking of objects.

In contrast to images, videos usually contain a lot of motion and various camera perspectives of objects. Hence, we can enhance the quantity of annotations when capturing a number of different views of an object. This raises the amount of available training data leading to a higher quality for further processing, especially for instance in machine learning. Inherently the effort of annotation may decrease.

2 System Description

The annotation tool consists of two components. The first one is the framework *Automated MOving Picture Annotator* (AMOPA) – a detailed description is provided by [18]. AMOPA integrates a huge set of concepts and process chains that are frequently used in image, video and audio processing. Access to video data is granted by the open source C-library FFMPEG, whereas this functionality is encapsulated and passed on to Java via *Streambaby* (cf. fig. 1). Recent extensions of the process concept of *Java Media Utility* (JMU) even allow the implementation of non-linear processing chains. By this, concepts of almost arbitrary topology can be realized. Relaying single image processing steps in the chain is achieved by a graphical editor or via XML. Every step is started as a single thread to fully utilize the capacity of multi-core CPU environments.

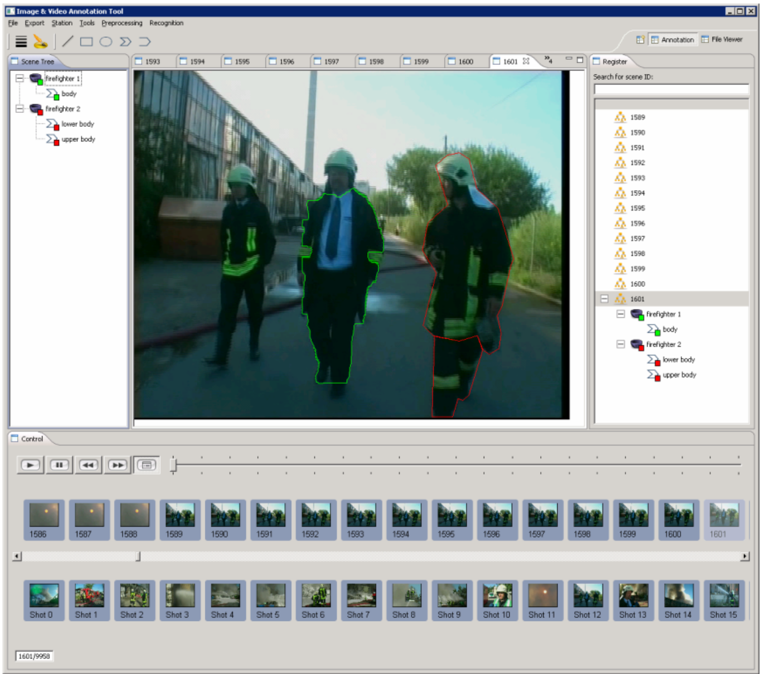


Fig. 2. Process of the annotation: a person is semi-automatically segmented (middle). The segmentation is show by a green polygon. The other person (right) is intellectually annotated by two red polygons indicating upper and lower body.

The second component basically comprises the annotation tool, which is developed in *Java RCP* and runs in Windows and Linux systems. The combination of *JFace* and *SWT* allows fast rendering of graphical content.

Fig. 2 shows the graphical user interface during the process of annotating. A classical menu bar provides functions to load, save, export data and to preprocess images and recognize objects. Preprocessing includes methods for shot boundary detection and interactive object segmentation. The recognition stage currently contains a method to track selected objects. The icon bar symbolizes options to create different figures like lines, rectangles, circles, and open and closed polygons as well as properties to set color and line width. The current annotation frame is shown in the center.

Object hierarchies are visualized by trees. The leftmost (scene) folder reflects a tree displaying the annotated content of the current frame. A more global content register containing all annotated frames of the video in form of trees is displayed on the right. Navigating through videos is realized by the control buttons below the annotation frame.

At the bottom of the screen two rows of images are presented. The upper row grants access to all frames of the current video. After the completion of the shot boundary detection, representative key frames of the corresponding shots are

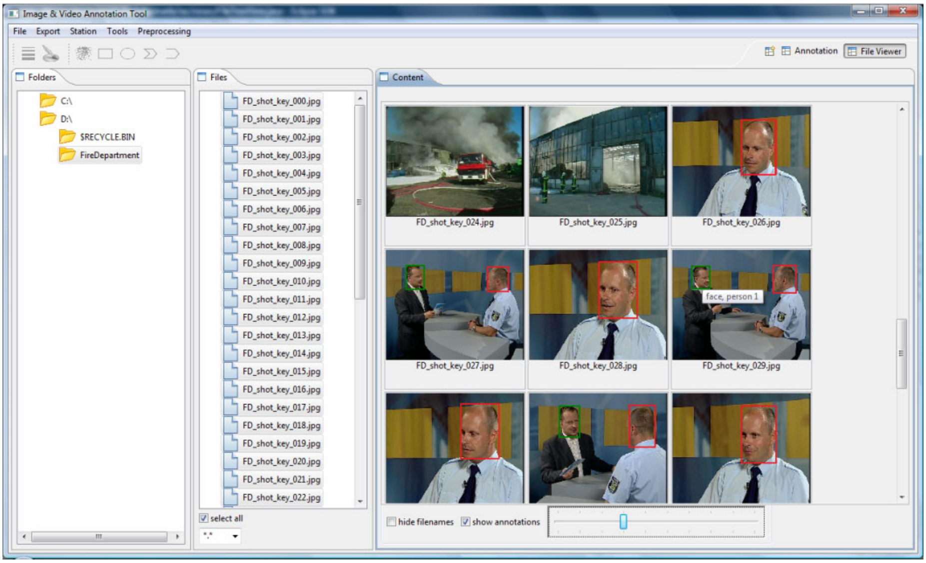


Fig. 3. The file viewer perspective of the annotation tool showing the annotated data as overlays on the image data

displayed in the lower row. Thus, the tool enables the user to scene-wise navigation as well as shot-wise navigation. Frames are transferred from the control to the annotation view via double clicking.

The latest feature of the annotation tool is a file viewer perspective that enables one to load and show images from disk in a scalable grid orienting on the well-known capabilities to view collections of thumbnails or photos within tools like *IrfanView* and *Microsoft Photo Gallery 2011*. Fig. 3 shows disk-stored representative key frames extracted from a fire department movie after the application of the shot detection. Two persons that occur in a dialogue scene have been consistently intellectually annotated with different colors (green and red) for each face over multiple images.

Microsoft Photo Gallery 2011 appends XML based annotations like the name and position of a face directly to image. In contrast our perspective loads annotated data from a separate file from the same folder and file name and overlays each image with its annotations. If the mouse is hovered over an annotation its description is shown in a toolbox tip.

As work in progress we currently prepare this perspective for (re)editing of overlays/annotations in the grid of the file viewer on the fly. Hence, simple annotation in multiple images can be done at once. On the other hand we target to implement the opportunity to rapidly evaluate the results of object classifiers on images or videos. Therefore, results must be stored within an compatible format to be visualized in the file viewer. An implemented logic function would allow to easily accept, reject or even score detections to grant further evaluations.

3 Applied Algorithms

In the following further insights into the available algorithms of the annotation tool are provided. Each method is implemented in a separate image processing chain within the framework AMOPA and can be controlled by an interface of shared objects.

3.1 Navigating in Videos and Data Export

Jumping to a specific frame within a video can turn out to be very difficult to implement, since the sequential streaming of the video has to be interrupted. Methods to solve this problem highly depend on the underlying video codec. This especially concerns the here applied version of FFMPEG when decoding MPEG video.

In contrast to decoding and object instantiation which is needed to finally access the data of a video frame, simple sequential decoding without object creation is multiple times faster. In simple benchmarks our system created between 70 and 100 accessible frames per seconds and decoded more than 1,200 frames per seconds. Therefore, currently we elude the problem of jumping by simply decoding the video data from the start to the requested time stamp.

When processing long videos, the performance of this method can be very poor because of longer reloading periods. Speeding up the process can eventually be achieved when indexing the *IFrames* of the video by assistance of an external tool like *Avidemux*[1]. Indexed access would allow us to jump directly onto the extracted byte position of the requested *IFrame* without any time loss.

Annotation data can be exported optionally to MPEG-7-DAVP [26], plain text or a customizable XML format.

3.2 Shot Boundary Detection

The detection of shot boundaries is based on [19], but actually uses only a small sample of the features proposed there to detect hard cuts (cf. [17]). Our test system (Dual Quad Core CPU, 3 GHz) achieved a processing speed of factor 2.2 times faster than real-time at 65 per cent workload in average, independent of the resolution of the video. Therefore, the video is first resized to half PAL resolution and then divided into disjoint blocks of *48 x 48* pixels.

Individual motion vectors are calculated for each block between two successive frames. The error between the original block and its motion compensation is computed by the minimum absolute distance of the blocks pixels. The ratio between the cumulative sum of the errors from all blocks and an additive function, smoothing the complete error sum of previous frames, triggers the detection of a transition.

---

[1] http://avidemux.org/

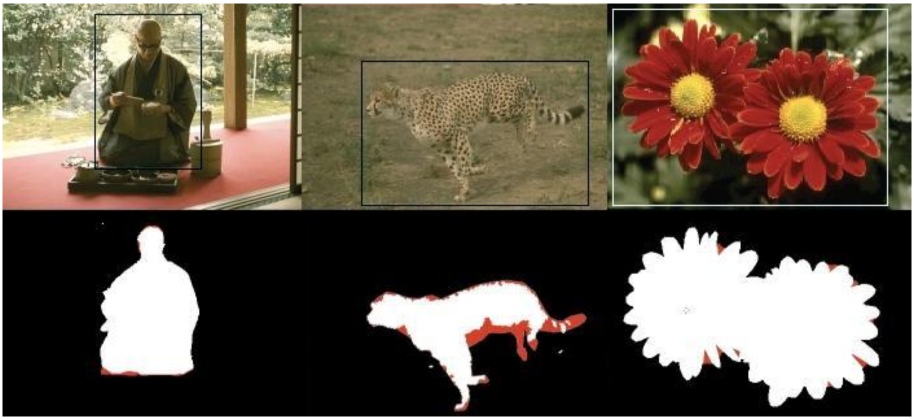


Fig. 4. Evaluation of the results of the *GrabCut* algorithm. Original images with selected regions in bounding boxes are displayed in the upper row. Grey color visualizes the differences to the manually created ground truth shown below. (modified from [24])

Table 1. Precision of the *GrabCut* method (information is given in pixels)

Image from Fig. 3	Overall area	GrabCut-Segmentation	Intellectual Segmentation	False detections	Error rate
Buddhist	151.526	24.826	24.501	575	2,35 %
Leopard	150.416	19.973	24.510	4.669	19,05 %
Flower	152.044	67.627	68.259	632	0,93 %

In contrast to the evaluation of [20], we were able to enhance the precision of the method by using MPEG-7 edge histograms within an environment of five adjacent frames of a shot candidate. Each candidate is tested for dissimilarity. Regarding the content of local television stations (> 100 hours of video), this method achieves detection rates of about 99 per cent with an empirical false-positive rate of max. 1.5 per cent.

3.3 Object Segmentation

Determining the exact contours of objects may be crucial for the process of feature extraction and subsequent processing. The intellectual annotation of such contours can be performed by free form or polygon tools, but is often time-consuming. To calculate the contours of an object, we propose the user to surround the object with a bounding box, before starting the semi-automated segmentation. Therefore, we use an implementation of [21], where two mixtures of Gaussians (MOG) are separately modeled for fore- and background. Pixels within the bounding box are assigned to the foreground, others to the background MOG. For each model the color reduction method of [22], which associates the pixels of the selected region to the most likely model is run. Then, its color distribution is computed. Afterwards, a graph is built by

Fig. 5. Tracking of the selected object (inner bounding box) within the initial image (left). Result of the automatic object tracking 24 frames later (right). Displayed imagery from TRECVID 2009 [1]. (modified from [24])

using the pixel distribution of both MOGs as source and sink to calculate the minimum cut [23]. This procedure is iterated until the set of pixels within both models remains constant. Results are visualized by fig. 4 and table 1. Acceptable results are achieved, if the object does not contain too much dissimilar colors. Problems also arise at strong textures, shadows as well as inhomogeneous background preventing a sharp distinction of the object.

3.4 Object Tracking

The annotation tool can track selected objects for a user defined number of frames. Fig. 5 outlines the process of tracking. First, an object is encircled by a bounding box, visualized by the blue inner rectangle on the left. The light yellow-green outer rectangle symbolizes the automatically calculated search window used for tracking the object within the next frame. Referring to [25], we apply a simple block-matching method with n step search by computing the minimum absolute distance from all pixels of the selected region and potential positions within the search window.

Optionally, the curved light blue line along the body of the walking person has been created by automated segmentation. In the current implementation this segmentation is optional and can be applied within every frame but does not affect the results of the tracking. Future work may aim to investigate methods to make this approach more reliable and even invariant to object deformation and changes. Usually, this can be achieved by the attempt to extract and track standardized object features from the bounding box or automated segmentation.

4 Future Work

We have described the current stage of development of an annotation tool that allows the intellectual annotation of semi-automated segmented and automatically tracked objects. Future implementations will concern functions to create ontologies and thesauri to make the tagging of objects more reliable and consistent. Furthermore, an object browser will support global searching for annotated objects within a video.

Another important step will be the integration of algorithms for the composition of shots to identify scenarios like dialogs and news broadcasting as well as to further enhance the exploited techniques for segmentation and tracking (cf. [27]). Methods for automated text and speaker recognition, and speech detection have already been developed within AMOPA. They are ready to be integrated into the annotation tool to enrich the set of functionality for annotation and analysis.

Acknowledgments. This work was partially accomplished within the project *sachsMedia – Cooperative Producing, Storage and Retrieval* (Project 03IP608), funded by the program Entrepreneurial Regions of the Federal Ministry of Education and Research, Germany.

References

1. Smeaton, A.F., Over, P., Kraaij, W.: Evaluation Campaigns and Trecvid. In: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval MIR 2006, New York, NY, USA, pp. 321–330 (2006)
2. Doermann, D., Mihalcik, D.: Tools and Techniques for Video Performance Evaluation. In: Proc. 15th International Conference on Pattern Recognition, vol. 4, pp. 167–170 (2000)
3. Lux, M.: Caliph & Emir: MPEG-7 Photo Annotation and Retrieval. In: Proceedings of the 17th ACM International Conference on Multimedia, Beijing, China, pp. 925–926 (2009)
4. Vezzani, R., Grana, C., Bulgarelli, D., Cucchiara, R.: A Semi-Automatic Video Annotation Tool with MPEG-7 Content Collections. In: Proceedings of the 8th IEEE International Symposium on Multimedia, San Diego, CA, USA (2006)
5. Petridis, K., Anastasopoulos, D., Saathoff, C., Timmermann, N., Kompatsiaris, Y., Staab, S.: M-OntoMat-Annotizer: Image Annotation. In: Linking Ontologies and Multimedia Low-Level Features. Engineered Applications of Semantic Web Session at the 10th Int'l. Conf. on Knowledge-Based & Intelligent Information & Engineering Systems, U.K. (2006)
6. Columbia University. IBM T. J. Watson Research Center: MARVEL: Multimedia Analysis and Retrieval System, `http://domino.research.ibm.com/comm/research_people.nsf/pages/jsmith.projects.html`
7. Naphade, M.R., Lin, C.-Y., Smith, J.R., Tseng, B., Basu, S.: Learning to Annotate Video Databases. In: Proc. SPIE, Storage and Retrieval for Media Databases vol. 4676, pp. 264–275 (2002)
8. Kounoudes, A., Tsapatsoulis, N., Theodosiou, Z., Milis, M.: A multi-level Video Annotation Tool based on XML-dictionaries. In: Proc. of the 10th WSEAS International Conference on Mathematical Methods, Computational Techniques and Intelligent Systems, Corfu, Greece (2008)

9. Luo, H., Eleftheriadis, A.: Designing an Interactive Tool for Video Object Segmentation and Annotation. In: Proc. of the 7th ACM International Conference on Multimedia, Orlando, FL, USA, pp. 265–269 (1999)
10. Goldman, D.B., Gonterman, C., Curless, B., Salesin, D., Seitz, S.M.: Video Object Annotation, Navigation, and Composition. In: Proc. of the 21st Annual ACM Symposium on User Interface Software and Technology, Monterey, CA, USA (2008)
11. Kipp, M.: Spatiotemporal Coding in ANVIL. In: Proc. of the 6th International Conference on Language Resources and Evaluation, LREC (2008)
12. Hagedorn, J., Hailpern, J., Karahalios, K.G.: VCode and VData: Illustrating a new Framework for Supporting the Video Annotation Workflow. In: AVI 2008, Napoli, Italy, pp. 317–321 (2008)
13. Schroeter, R., Hunter, J., Kosovic, D.: FilmEd: Collaborative Video Indexing, Annotation and Discussion Tools Over Broadband Networks. In: Proc. of the 10th IEEE International Conference on Multimedia Modeling, Los Alamitos, California, USA, pp. 346–353 (2004)
14. Concejero, P., Munuera, J., Lorenz, M.: The MESH Mobile Video Annotation Tool. In: Proc. of the 5th ACM Nordic Conference on Human-computer Interaction: Building Bridges, NordiCHI 2008, Schweden (2008)
15. Ritter, M., Eibl, M.: Ein erweiterbares Tool zur Annotation von Videos. In: Proc. 12th International Symposium on Information Science 2010, Hildesheim, Germany (2010) (in press)
16. Fan, J., Gao, Y., Hangzai, L., Jain, R.: Mining Multilevel Image Semantics via Hierarchical Classification. IEEE Transactions on Multimedia 10(2), 167–187 (2008)
17. Ritter, M., Eibl, M.: Visualizing steps for shot detection. In: LWA 2009: Lernen - Wissen - Adaption, Workshop Proceedings, Darmstadt, Germany, pp. 98–100 (2009)
18. Ritter, M.: Visualisierung von Prozessketten zur Shot Detection. In: Workshop Audiovisuelle Medien: WAM 2009, Technical Reports of Computer Science (CSR-09-04), Chemnitz University of Technology, Germany, pp. 135–150 (2009)
19. Liu, Z., Gibbon, D., Zavesky, E., Shahraray, B., Haffner, P.: AT&T RESEARCH AT TRECVID 2006 Workshop Contribution. AT&T Labs-Research (2006)
20. Zwicklbauer, S.: Evaluierung und Implementierung von Shot-Boundary-Detection-Algorithmen zur automatischen Video-Anno¬tation. Bachelor thesis, Universität Passau, pp. 48–52 (2010)
21. Talbot, J.F., Xu, X.: Implementing GrabCut. Brigham Young University, Provo, UT, USA (2006), `http://www.justintalbot.com/course-work/`
22. Orchard, M., Bouman, C.: Color Quantization of Images. IEEE Transactions on Signal Processing 39(12), 2677–2690 (1991)
23. Boykov, Y., Kolmogorov, V.: An Experimental Comparison of Min-cut/Max-flow Algorithms for Energy Minimization in Vision. IEEE Trans. on Pattern Analysis and Machine Intelligence 26(9), 1124–1137 (2004)
24. Höhlig, S.: Analyse und Implementierung eines Verfahrens zur interaktiven semi-automatischen Objektmarkierung und -verfolgung. Bachelor thesis, Chemnitz University of Technology (2010)
25. Beck, P.: Implementierung eines flexiblen Algorithmus zum Tracking von Objekten in Java. Seminar paper, TU Darmstadt, Germany (1999)
26. Bailer, W., Schallauer, P., Neuschmied, H.: Description of the MPEG-7 Detailed Audiovisual Profile (DAVP). Technical Report, Joanneum Research, Graz, Austria (2007)
27. Price, B.L., Morse, B.S., Cohen, S.: LIVEcut: Learning-based Interactive Video Segmentation by Evaluation of Multiple Propagated Cues. In: Proc. International Conference on Computer Vision (ICCV 2009), Kyoto, Japan (2009)

Development of an Augmented Feedback Application to Support Motor Learning after Stroke: Requirement Analysis

Simone Schmid*

Berlin Institute of Technology, Center of Human-Machine Systems,
Research Training Group prometei,
Franklinstr. 28-29, 10587 Berlin, Germany

Abstract. The introduction of robotic and mechatronic devices in motor rehabilitation after stroke offers new possibilities to support the relearning of motor skills. One of these new fields in rehabilitation research is the visual augmentation of movement-related parameters. So far, the effects of augmented feedback on the learning outcome are only vaguely assessed and more clinical studies are needed. This paper describes the first step of a user-centered approach to develop visually augmented feedback applications for the use in clinical studies.

Keywords: augmented feedback, motor learning, neurorehabilitation, stroke, requirement analysis, human-centered design.

1 Introduction

Stroke is the leading cause of adult disability [27]. Each year, approximately 780,000 individuals in the United States and 250,000 in Germany suffer from the aftermath of stroke [22].The most prevalent consequences of stroke are motor deficits like hemiparesis. Hemiparesis is amongst others characterized by abnormal muscle tone, postural adjustments and movement synergies, incorrect timing within movement patterns, loss of inter-joint coordination, and loss of sensation [6]. These impairments affect the patients' quality of life and their ability to perform activities of daily living. The prospect of returning to an independent life after stroke can be increased by an early motor therapy. The recovery of motor functions is a drawn-out process based on intensive, task-specific and repetitive movement training, supported by manual as well as device-assisted therapy [20]. This project focuses on the recovery of the upper extremity (UE) because arm and hand dysfunctions persist for many years after the stroke [22].

The introduction of device-assisted motor rehabilitation allows recording, analyzing, extracting and presenting data about the current movement on- and off-line. Information about pathological movement patterns, deviations of trajectories and/or force profiles can be quantified and automatically fed back to the

* This work was supported by the Research Training Group *prometei* at TU Berlin funded by the German National Science Foundation (DFG).

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 305–314, 2011.
© Springer-Verlag Berlin Heidelberg 2011

patients. The presented movement-related parameters, also referred to as augmented feedback, could serve as a basis for error correction and thereby facilitate motor learning [15].

So far, effects of augmented, visual feedback on re-learning of motor skills after stroke have not been elaborated. Before any further investigations can take place, feedback visualizations must be developed that can be perceived and processed by the patients. This is especially true because 56.6% of the stroke survivors are seriously affected by cognitive deficits such as disturbances of attention and executive dysfunction [23],[12].

This paper describes the user-centered development approach of the graphical feedback application tailored to the needs of stroke patients. For that, the importance of augmented feedback for motor learning after stroke is outlined. Results of the requirement analysis are reported and conclusions in the form of project-specific usability goals and design principles are presented.

2 Relevance of Augmented Feedback in Motor Rehabilitation after Stroke

Besides practice, augmented feedback is one of the most powerful factors identified to influence motor learning in healthy individuals [15],[16]. In contrast to inherent feedback that naturally results from the execution of movements, augmented feedback is provided by an an external source. Inherent feedback is considered as complex and hard to interpret, whereas augmented feedback can emphasize important movement aspects individually [24]. There are different facets of augmented feedback such as modality, frequency, scheduling and content affecting the learning outcome dependent on the complexity of the movement to be learned, the learners skill level and the context.

The same influencing variables on motor learning in healthy individuals and stroke patients can be supposed due to the assumption of similar learning mechanisms [29]. For example, feedback can be used to reduce movement errors because it shapes neural activity in motor and premotor areas [1]. Single unit recordings demonstrate that visual information can have effects on neural circuits and therefore provide a potent signal for reorganization of sensorimotor circuits [13]. Some researches suggest that stroke patients could even benefit more from adequately presented augmented feedback than healthy individuals [9]. They reason that stroke patients are not able to use intrinsic or proprioceptive feedback because of the cognitive, perceptual and sensation impairments and need supplementary information.

There are many systematic studies about the effects and mechanisms of feedback in healthy individuals. The increasing number of reviews about augmented feedback after stroke published in the last few years reflects the importance of feedback in neurorehabilitation (see i.a. [28],[7],[26]). Besides, the reviews indicate that the used materials for feedback presentation are neither standardized nor guarantee the suitability for stroke patients.

3 Development Process

The development of the graphical feedback application follows the rationale of user-centered design (UCD). UCD focuses on the user demands through planning, design and development. For this, findings from the field of software ergonomics, in turn obtained from disciplines like computer sciences, linguistic, psychology and sociology are adapted to the particular needs of the users. The core objective is to design easy to use products, allowing an optimal interaction according to a pleasurable user experience [18].

One approach to UCD is usability engineering which provides explicit methods to develop software that corresponds to usability principles [8]. Usability engineering offers various process models like the Usability Engineering Lifecycle [18]. Figure 1 illustrates the procedure applied in this project which is mainly based on the models described in [5] and [18].

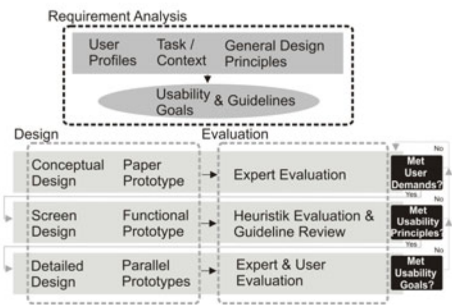


Fig. 1. Two-phased process model underlying the development of the feedback application. It consists of a requirement analysis followed by several levels of design and evaluation (adapted from [5] and [18]).

The development process is comprised of two parts. In the requirement analysis the users will be identified and relevant characteristics will be assessed. In the task and environment analysis the user's major tasks, goals end information needs will be determined. General design principles will be reviewed and principles relevant to the project will be extracted. The obtained information will lead to high-level usability goals driving further design decisions. They will be transfered to more detailed design guidelines that in turn will be specified in project standards. These standards are directly applicable and will be summarized in styleguides. The formulation of a styleguide is an ongoing process requiring refinement as the project progresses.

Based on the results of the requirement analysis the design phase will be entered. It consists of three design levels, each associated with an evaluation action. In the first level of design, major display structures and information visualization ideas will be generated and implemented in a paper prototype. If all respondents agree upon the concept, a screen design prototype will be produced and evaluated by a combination of a heuristic evaluation and guideline review based on the usability principles. After a refinement and reevaluation a detailed user interface design will be developed and tested by the users.

4 Requirement Analysis

4.1 General Considerations

The clinical context imposes some major restrictions concerning the study setting and user group. For the clinical study, two rehabilitation devices will be used: the arm trainer Bi-Manu-Track (BMT) and the Reha-Slide (RS)(see figure 2). The BMT is a microprocessor-controlled device for reproducing active and passive flexion and extension of the wrist as well as pro- and supination of the forearm movements. The RS is used for the treatment of arm, shoulder and elbow by flexion and extension of the elbow, abduction and adduction of the shoulder and rotation of the wrist. (For a comprehensive description of the BMT and RS see [10], [11].)

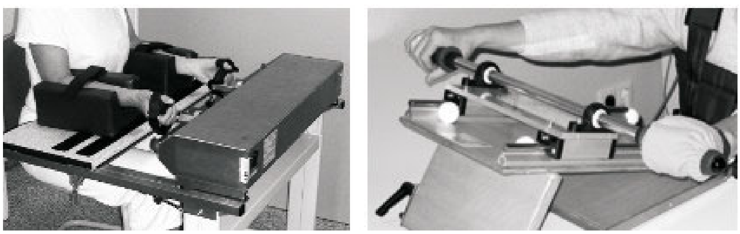


Fig. 2. Bi-Manu-Track (left), Reha-Slide (right)

Only patients who meet certain in- and exclusion criteria are eligible to participate in the device-assisted therapy and clinical study. According to [2] these criteria are: a) minimally functional UE, b) no severe upper limb spasticity, c) no hemiparetic shoulder pain, d) no swollen hand impeding closing the fist, e) no other neurological or orthopedic impairments of the UE, f) mobilized in a wheelchair, g) able to follow therapists instructions, h) a positive single test session in the arm studio, and i) informed consent to participate.

4.2 Designs, Materials, and Procedures

User Profiles. The first step in generating user profiles is to define the intended users. This is already done by the criteria described above. However, the patients could significantly differ in characteristics relevant to the user interface design such as preferences, attitudes, skills, knowledge and experience. Representative data about user characteristics are usually collected with user profile questionnaires. In the clinical setting questionnaires are not suitable because most of the patients are not able to read and to fill out the form independently due to their cognitive and motor limitations. Therefore, structured interviews based on typical user profile questionnaires are deemed as the best method to get direct, accurate and reliable information. In structured interviews the sequence and wording of each question is fixed providing consistency across all interviewees. Besides, predefined response categories facilitate the formulation of answers. The direct conversation allows the interviewer to be sensitive to misunderstandings or difficulties in speech production.

The used interview guide consists of questions about demography (age, family status, educational background), medical background (previous and current physical or mental dysfunctions, etc.) as well as computer usage (availability of computers, frequency of usage, etc.) and experience (interest in computers, preferences, etc.). Additionally, some adapted items from the TA-EG, a questionnaire about affinity for technical products [14] are included.

The patients were recruited at the arm studio of a rehabilitation center to ensure that they meet the above-noted criteria. The arm studio contains several arm rehabilitation devices to intensify the UE treatment. (For a detailed description of the arm studio see [2].) After getting the approval of the rehabilitation center, the interviews were conducted by two different but trained interviewers. The interview began with a written overall statement of purpose. A cover letter and informed consent was handed to the patients. The patients signed the informed consent and by that declared to voluntarily participate. The interviewers read out load the questions and the patients were free to read along. The interview took on average 30 minutes.

The answers were recorded on the interview guide and immediately classified into the predefined response categories. The data were analyzed according to commonalities of and differences between the users.

Task and Environment Analysis. The task and environment analysis serves as a basis for developing the movement task and strives for a better understanding of the information needs of the patients. Furthermore, information about the therapy schedule to assure a smooth integration of the study session into the tight time plan of the patients can be gained. Implications about general conditions of the study and the detailed study procedure will not be reported here.

The task and environment analysis was conducted by a semi-structured observation. This approach allows the concentration of actions that have been considered as relevant in advance, without missing important events that arise out of the situation. The observation took place in the arm studio. The recruitment of the patients corresponded the procedure described in the user profile section. In case some questions arise, the observations could have been discussed with the mentoring therapist in the arm studio. No videotaping was allowed to avoid the patients feeling uncomfortable.

The observer was instructed to list every action and event in a serial manner and was trained to pay attention to occurring problems and disturbances. The observer should also code all events according to the structured observation protocol. Categories in the observation protocol are related to the physical environment, distractions and interruptions as well as events and facts concerning the task. These predefined categories were also used for the data analysis.

General Design Principles. General design principles are fundamentals reflecting knowledge about human information processing and perception [17]. These principles can be used for evaluations in the design and evaluation phase. Project-relevant principles were identified by reviewing commonly cited usability literature like [19],[25], and [21], as well as standards like [4],[3], and [5]. These general principles were translated to project needs based on the specification from the user profiles as well as the task and environment analysis.

5 Results

5.1 User Profiles, Task and Environment Analysis

12 patients (50% female) participated in the user profile interview ranging in age from 21 to 76 years. Their mean hospitalization was 4 weeks. 41.7% are married and living with their partners. The educational level is quite high with 58.3% of the patients having an Abitur or higher. About 75% of the patients wear corrective lenses and/or suffer from hearing disorders. They also have perceived attentional deficits and speech production difficulties. 91% could imagine that computers could help and support the motor rehabilitation.

The patients widely differ in their attitudes towards and experience with computers. Two typical user groups could have been identified: experienced and inexperienced users. Prototypical experienced users own a computer and use it on a regular basis at least once a week. They are familiar with word processing and internet. These users are comfortable with the keyboard and computer mouse. They are generally positively attuned to computers and they like learning new computer programs. On the other hand, inexperienced users possess no computer at home, some of them never have used a computer once in their live. They avoid using computers in general and could not imagine to feel comfortable with either of the input devices. They know computer-assisted learning software from their neuropsychological therapy sessions but they are skeptical about the advantages. They are not interested in anything concerning computers and some are even afraid of them. They are also convinced that they will never learn how to deal with computers and computer programs.

The task analysis suggests that there are three main training objectives: extending the range of motion, controlling the motion according to steadiness, a combination of range of motion and steadiness. For the development of the first feedback application, the project team and clinicians agreed to concentrate on the range-of-motion task. A typical task formulation for the BMT could be "Increase your range of motion. Therefore, turn your wrist as far as possible to the right and left." Technically, there are several parameters that can be recorded and visualized such as the maximum range of motion patients are able to reach, current movement, the maximum achieved range of motion, the range of motion achieved in the last trial, etc. The decision about the parameters that will be recorded and visualized will be made in the design phase.

The environment analysis showed that the patients are scheduled to their treatments but arrive with an average delay of 10 minutes. The therapist prepares the rehabilitation devices and seats the patients in front of them. The patients absolve their training sessions without any assistance from the therapist. The sessions last a averaged period of 15 to 20 minutes and a total of 100 to 300 repetitions. The therapist returns as soon as the training routine is completed. The task is quite monotone but mentally demanding. Most of the patients are highly distractible and confused as well as easily exhausted and fatigued. Frequently, the therapist has to remind patients to stay awake or to refocus their concentration on the movement task.

5.2 Usability Goals

The user profile analysis confirmed previous findings about the prevalence of cognitive deficits affecting the information processing capacity and speed of the patients. Therefore, the feedback application should be supporting rather than further strain the cognitive system. The identification of inexperienced and apprehensive users indicates that the application must be convincing of being helpful and supporting. Thus, the design must be motivating, easy to understand and to use. Most of the patients have some kind of visual disorder and difficulties with visual thinking. Hence, the application must consist of easy to perceive and visible elements without being demanding to the visual system. The monotone movement task and the high distraction sensitivity as well as the rapid tiredness require the adoption of a fine balance between a stimulating and attention attracting application and prevention of interference with the perception and execution of the actual movement. Due to the frequent use, the application should not contain annoying and reoccurring instructions. The usability goals can be summarized as follows:

– The design must support learning without being mentally demanding.
– The design must support inexperienced, anxious users.
– The design must support perceptibility and visibility.
– The design must attract the attention without being interfering.
– The application should support frequent use.

Even if the usability goals help focus the design efforts, they are not measurable and no concrete design alternatives or ideas can be inferred. Based on the remarks above, the usability goals can be broken down in more detailed design principles. These project-specific principles can be amplified with suggestions of the general design principles from the literature review. The key principles identified so far are summarized in table 1.

The principles will be used for heuristic evaluations and as criteria for user tests. In the course of the development the guidelines will be specified and transfered to project standards. As explained above, the standards will be available

Table 1. Conclusions about major design principles and exemplary guideline suggestions according to the usability goals derived from the requirement analysis

Principle	Description	Guideline
Suitable for perception	Elements and objects of the interface must be visible, distinguishable and readable.	Use high color contrasts, big font sizes, readable font types, ...
Easy to understand	The application is comprehensible if it is self-descriptive, relevant, consistently and clearly structured and conform with users expectations.	Use familiar and self-descriptive symbols and expressions, consistent colors, shapes and positioning, ...
Suitable for the task	The application is suitable for the task if it supports the movement execution and learning process effectively and efficiently.	Provide compatibility of input (movement) and output (feedback), present only task relevant information, visualize the information compactly, ...
Suitable for motivation	The application must be stimulating and interesting instead of frustrating and boring.	Chose parameter values that are encouraging and make improvements visible, use friendly and pleasant colors and graphics, ...

in a styleguide and can be used for similar applications. For instance the requirement of being suitable for motivation results in the guideline suggesting to use friendly and pleasant colors. The standards could contain specifications about the background color fixed to a concrete hex color value.

6 Outlook

This paper suggested to apply a user-centered design approach to develop a user interface for the presentation of movement-related parameters. The application will be used in a clinical feedback study that is aimed to investigate effects and mechanisms of visual, augmented feedback to support the motor learning process after stroke. It was expected of the user-centered design approach to produce an application which is tailored to the special needs of the stroke patients concerning different cognitive and motivational conditions as well as computer attitudes and experiences.

As reported, the first phase of the development process described in figure 1 has been successfully completed. Usability goals, principles and guidelines have been derived. Up to date, the level of conceptional design and expert evaluation as well as the screen design level with the heuristic evaluation and guideline reviewing is completed and the detailed design level has been entered. Currently, work on alternative designs of the application is in progress based on the suggestions of the preceding design levels and evaluations.

The retrieved results and undergone experiences are very encouraging. The adaption of the user-centered approach and the extensive requirement analysis as well as evaluation actions after every design level are paying off. Especially the cognitive, perceptual and motor deficits of the stroke patients made an early involvement of the users essential and fruitful. Many important suggestions concerning the design could have been obtained from the users and clinicians. In addition, the intended use of the application in a clinical study necessitates the approval of the clinicians. Therefore, the user-centered approach is a excellent instrument to increase the acceptance amongst the stakeholders.

References

1. Bray, S., Shimojo, S., ODoherty, J.P.: Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. The Journal of Neuroscience 27(28), 7498–7507 (2007)
2. Buschfort, R., Brocke, J., Heß, A., Werner, C., Waldner, A., Hesse, S.: Arm studio to intensify upper limb rehabilitation after stroke: Concept, acceptance, utilization and preliminary clinical results. Jounral of Rehabilitation Medicine 42, 310–314 (2010)
3. CEN: DIN EN ISO 9241-12: Ergonomic requirements for office work with visual display terminals – Part 12: Presentation of information. Beuth Verlag (1998)
4. CEN: DIN EN ISO 9241-110: Ergonomics of human-system interaction- Part 110: Dialogue principles. Beuth Verlag (2006)
5. CEN: DIN EN ISO 9241-210: Ergonomics of human-system interaction – Part 210: Human-centered design for interactive systems. Beuth Verlag (2010)
6. Cirstea, M.C., Levin, M.F.: Compensatory strategies for reaching in stroke. Brain 123, 940–953 (2000)
7. van Dijk, H., Jannink, M.J.A., Hermens, H.J.: Effect of augmented feedback on motor function of the affected upper extremity in rehabilitation patients: A systematic review of randomized controlled trials. Journal of Rehabilitation Medicine 37(4), 202–211 (2005)
8. Dix, A., Finley, J., Abowd, G.D., Beale, R.: Human-computer interaction, 3rd edn. Pearson, New York (2004)
9. Flinn, N.A., Radomsky, M.V.: Occupational therapy for physical dysfunctions. In: Trombly, C.A., Radomsky, M.V. (eds.) Occupational therapy for physical dysfunction, 6th edn., ch.14, pp. 283–297. Lippincott Williams & Wilkins, Baltimore (2006)
10. Hesse, S., Schulte-Tigges, G., Konrad, M., Bardeleben, A., Werner, C.: Robot-assisted arm trainer for the passive and active practice of bilateral forearm and wrist movement in hemiparetic subjects. Archives of Physical Medicine and Rehabilitation 84(6), 915–920 (2003)
11. Hesse, S., Werner, C., Pohl, M., Mehrholz, J., Puzich, U., Krebs, H.I.: Mechanical arm trainer for the treatment of severely affected arm after stroke: A single-blinded randomized trial in two centres. American Journal of Physical Medcine Rehabilitation 87(10), 779–788 (2008)
12. Jokinen, H., Kalska, H., Mäntylä, R., Pohjasvaara, T., Ylikoski, R., Hietanen, M., Salonen, O., Kaste, M., Erkinjuntti, T.: Cognitive profile of subcortical ischaemic vascular disease. Journal of Neurology, Neurosurgery, and Psychiatry 77(1), 28–33 (2006)
13. Kakei, S., Hoffman, D.S., Strick, P.L.: Sensorimotor transformations in cortical motor areas. Neuroscience Research 46, 1–10 (2003)

14. Karrer, K., Glaser, C., Clemens, C., Bruder, C.: Technikaffinität erfassen - der Fragebogen TA-EG. In: Lichtenstein, A., Stoessel, C., Clemens, C. (eds.) Der Mensch als Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme (ZMMS Spektrum, Reihe 22, Nr. 29), pp. 20–196. VDI Verlag GmbH, Düsseldorf (2009)
15. Magill, R.A.: Motor learning and control - Concepts and applications, 7th edn. McGraw-Hill, New York (2007)
16. Marschall, F., Daugs, R.: Feedback. In: Mechling, H., Munzert, J. (eds.) Handbuch Bewegungswissenschaft - Bewegungslehre, ch. D.2, pp. 281–294. Hofmann, Schorndorf (2003)
17. Mayhew, D.: A design process for web usability. In: Proctor, R.W., Vu, K.-P.L. (eds.) The Handbook of Human Factors in Web Design, ch.19, pp. 338–356. Lawrence Erlbaum Associates, Mahwah (2005)
18. Mayhew, D.J.: The usability engineering lifecycle: A practitioner's handbook for for user interface design. Morgan Kaufmann, San Francisco (1999)
19. Nielsen, J.: Usability engineering. Morgan Kaufmann, San Francisco (1993)
20. Pohl, M., Werner, C., Holzgraefe, M., Kroczek, G., Mehrholz, J., Wingendorf, I.: Repetitive locomotor training and physiotherapy improve walking and basic activities of daily living after stroke: A single-blind, randomised multi-centre trial. Clinical Rehabilitation 1, 17–27 (2007)
21. Preiser, W., Smith, K.H.: Handbook of universal design, 2nd edn. McGraw-Hill, New York (2010)
22. Rosamond, W., Flegal, K., et al.: Heart disease and stroke statistics-2008 update: A report from the American Heart Association statistics committee and stroke statistics subcommittee. Circulation 117, e25–e146(2008)
23. Saxena, S.K.: Prevalence and correlates of cognitive impairment in stroke patients in a rehabilitation setting. International Journal of Psychosocial Rehabilitation 10(2), 37–47 (2006)
24. Schmidt, R.A., Lee, T.D.: Motor control and learning. Human Kinetics, Champaign/Ill (2005)
25. Shneiderman, B., Plaisant, C.: Designing the user interface: Strategies for effective human-computer interaction, 5th edn. Addison-Wesley Longman Publishing Co., Boston (2010)
26. Subramanian, S.K., Massie, C.L., Malcolm, M.P., Levin, M.F.: Does provision of extrinsic feedback result in improved motor learning in the upper limb poststroke? A systematic review of the evidence. Neurorehabilitation and Neural Repair 24(2), 113–124 (2010)
27. Truelsen, R.B., Mathers, C., Boudousslavsky, J., Boysen, G., Piechowski-Józwiak, B.: Stroke incidence and prevalence in Europe: A review of available data. European Journal of Neurology 13, 581–598 (2006)
28. van Vliet, P.M., Wulf, G.: Extrinsic feedback for motor learning after stroke: what is the evidence? Disability and rehabilitation 28(13-14), 831–840 (2006)
29. Winstein, C.J., Merians, A.S., Sullivan, K.J.: Motor learning after unilateral brain damage. Neuropsychologica 37, 975–987 (1999)

2D vs. 3D Pain Visualization: User Preferences in a Spinal Cord Injury Cohort

Fotios Spyridonis and Gheorghita Ghinea

Department of Information Systems and Computing, Brunel University, Uxbridge, UK
{Fotios.Spyridonis,George.Ghinea}@brunel.ac.uk

Abstract. Research on pain experienced after Spinal Cord Injury (SCI) has revealed that not only are there several types of pain present in the same individual with this kind of trauma, but also that people who suffer such an injury can describe the characteristics of the same type of pain in different ways. Making it possible, therefore, to more precisely describe pain experience could prove to be vital for an increased quality of life. Accordingly, fifteen individuals with pain after SCI were asked to describe their pain experience using a 3 Dimensional (3D) model of the human body that could be used as an aid in communicating their pain. The results of this study suggest that the consensus of the participants approved the ability of the 3D model to more accurately describe their pain, an encouraging outcome towards the use of 3D technology in support of post SCI pain rehabilitation.

Keywords: Rehabilitation, Chronic Pain, Spinal Cord Injury, 3D Visualization, Pain Drawing.

1 Introduction

Studies on the implications of Spinal Cord Injury on patients have exposed that post traumatic pain is amongst its most prevalent consequences, coming only second after loss of function. Reports from individuals who have suffered this kind of trauma reveal that the severity of such pain experience could be so immense that it often interferes with sleep and everyday activities [4]. In fact, a summary of results from several studies [see 16, 19] indicates that the average reported estimate of the prevalence of chronic SCI pain is approximately 65%, with roughly one-third of those affected reporting the severity as greater than 7 in a scale of 10 on a Visual Analogue Scale (VAS). This revelation could justify the need of the nearly 40% of SCI patients to trade their pain with additional loss of mobility, bladder or bowel control, and sexual function [16].

As a result of its incidence, considerable research efforts towards pain relief have been reported so far [7, 9]. Nevertheless, the majority of them have been criticized in the clinical literature regarding their applicability in assessing persons with SCI. The reason behind this controversy lays in the multidimensional nature of pain, which is characterized by physical discomfort, and is often influenced by complex qualities associated with psychological and cultural factors [3, 10, 13]. Owing to its subjectivity, it is therefore argued, that individuals who have to deal with pain after

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 315–322, 2011.
© Springer-Verlag Berlin Heidelberg 2011

SCI may frequently experience substantial difficulties when it comes to precisely describe their pain characteristics, as they may have been influenced by the above factors, resulting in different interpretations of the same pain experience.

To this end, considering the heterogeneity of the pain experience in this patient population, enabling the individual to also visually communicate his/her pain was further introduced as a supplementary approach for the intended purpose. Accordingly, the focus of the work described in this paper has been to examine from the user perspective the applicability of a 3D visualization approach in the communication of pain, in the anticipation that it will constitute a significant improvement over the current approaches and make an important contribution towards the effective assessment of pain for persons with SCI. Subsequently, the structure of this paper is as follows: Section 2 presents an overview of the current methods used for pain visualization, while section 3 discusses the methodology used to evaluate our 3D model as compared to the current approaches. Finally, section 4 presents the results of our study, and section 5 concludes with our study's discussion.

2 Current Approaches to Pain Visualization

The application of 2-Dimensional (2D) visualization techniques in the assessment of pain is not a recent trend. Over the last decades, a simple self-assessment method – the 'pain drawing' (Fig. 1) – has been exponentially applied across several medical conditions in the attempt to assist patients to provide their pain information. Originally proposed in the 1940s by [12, 14], this technique is considered to be a popular visual aid tool among the clinical field, as it enables patients to describe, using a monochrome symbol, the spatial location, as well as the type (*usually ache, pain, pins and needles, and numbness*) of their pain on a 2D human diagram. Considering its wide applicability, a number of research studies have been conducted over the years to examine its validity for the intended purpose. Their results indicate that the pain drawing is considered to be a valuable and useful tool in identifying pain location and sensation type, with most of patients demonstrating consistency in completing it [12, 18].

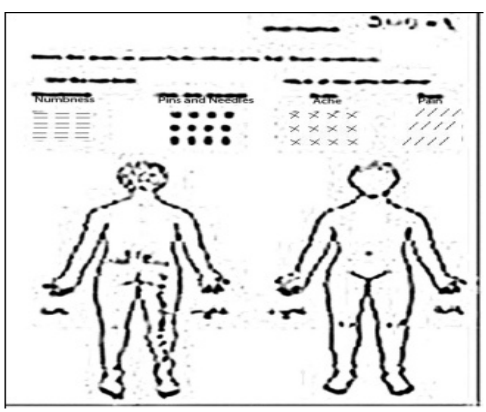


Fig. 1. 2D Pain Drawing

Due to their established usefulness and acceptance in assessing pain across various medical conditions [8, 11, 12], there have also been efforts to utilise the benefits of pain drawings to the SCI population that is characterised by the severity of this injury, which appears, for instance, in approximately more than 1,000 people per year in Britain alone[17].

To this end, [15] have used pain drawings to assess back pain and spinal deformity, whereas [4] attempted to identify and relieve the most disturbing pains in patients with a SCI that affect the quality of their life. Similarly, [5] utilised pain drawings before and after surgery to evaluate treatment interventions in SCIs.

2.1 Exploring the Need for 3D in the Assessment of SCI Pain

This situation, however, is changing rapidly. Advances in computer graphics technology towards a 3-Dimensional perception of the environment offer the ability for a more accurate and interactive experience. As a result, recent technological advancements have enabled the use of such 3D visualization expertise in offering promising opportunities for clinical rehabilitation. For instance, the use of 3D interactive representations of the head and arm has shown to be beneficial for patients in terms of enhancing their motor performance after a stroke [1], while innovations in 3D techniques have enabled the reconstruction of facial prosthesis for cancer patients [2]. The visualization capabilities that 3D technology provides could be similarly exploited towards the design of novel assistive rehabilitation solutions for persons with pain after Spinal Cord Injury.

To this end, after consultations with medical staff in the Royal National Orthopaedic Hospital in London, UK, it was considered more beneficial that patients would be offered the ability to describe the pain that they were experiencing onto a 3D model of themselves rather than a 2D one, as the former provides a continuous body surface, which allows patients to better map their pain. As a result, by expanding on past research, an adaptation of the 3D pain drawing (Fig. 2b) that was initially devised by [6] (Fig. 2a) was developed. This revised 3D model was considered to be more suitable for the purpose intended, in the anticipation that it would remove any anomalies caused by the limited depth perception of the 2D pain drawing currently used in SCI pain assessment, while addressing statements of the form *"I have a pain on the inside of my thigh",* which are not easily captured in a 2D visualization of pain.

Accordingly, while the concept of 2D drawings is well-established amongst the pain community, the usefulness of their 3D equivalent is yet to be examined. In this paper, therefore, we report the results of an empirical study, which aimed to *identify the user preferences regarding the most well perceived way of visualizing their pain, as derived by comparing both the 2D and the 3D methods while used in a natural clinical setting*.

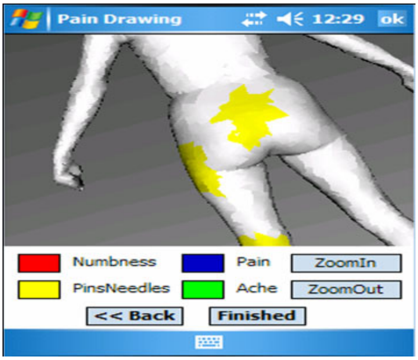


Fig. 2a. Initial 3D Pain Drawing

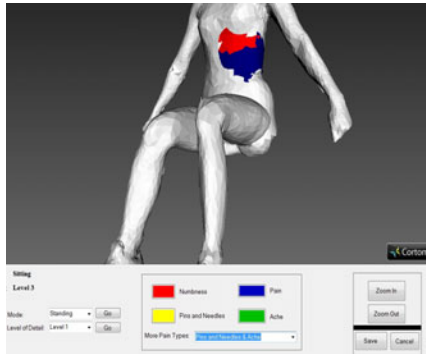


Fig. 2b. 3D Pain Drawing for SCIs

3 Methodology

Consequently, the objectives of this study are: *a. to investigate the feasibility of the 3D model in visualizing patient SCI pain characteristics*, as well *as b. to examine its usability in achieving the above*. To keep with best practice, our developed 3D model will be evaluated against the well-established 2D pain drawing of Fig. 1. Therefore, to address our research objectives, each participant will be given the chance to use both methods and at the end fill an evaluation questionnaire about the feasibility and usability of using these two methods for the purpose of recording and visualizing their pain experience.

3.1 Instrumentation and Materials

The instrumentation used for this study consists of a laptop that runs the 3D pain drawing, the paper-based 2D pain drawing, and an evaluation questionnaire. The aforementioned questionnaire is an evaluation survey for both the 2D and 3D pain drawings, in which patients are asked to record their opinions about both methods on a Likert scale of 1 (Disagree) to 5 (Agree) (Table 1 below).

Table 1. Participant Evaluation Questionnaire

Questions
Q1. It was easy to log pain information on the pain diagram
Q2. Showing the exact type and location of my pain on the pain drawing was easy
Q3. I believe the pain drawing was insufficient to express my pain
Q4. How would you describe the overall layout of the interface?

3.2 Description of Subject Group

The research sample consisted of 15 persons with SCI (7 female; 8 male, mean age 52.3 years, range 28-75) who volunteered to participate in the research study between July 2010 and October 2010. All were recruited from the Spinal Cord Injury Unit in the Royal National Orthopedic Hospital, London, UK. The diagnosis varied, but the majority was identified with traumatic SCI. The criteria for selection was that the participant has spinal cord trauma, has an age of 20 years or greater and experiences some pain during the period of study. Finally, the range of pain intensity varied from 0-9, with the mean maximum pain intensity being 8.375 on a Visual Analogue Scale (VAS).

3.3 Evaluation Protocol

After consultations with the clinicians, it was decided that the data collection would take place in four points in time over a period of one day for each participant, with an approximately 2-3 hour time difference between these four recordings (between 8.30am and 5pm) in order to address the change of pain over time. For the first two recordings, the participant was asked to report any pain experience on the paper-based, 2D pain drawing, whereas for the remaining 2 recordings s/he was similarly asked to report any pain on the laptop-based, 3D pain drawing provided. Each session lasted approximately 20 minutes. At the end of the first two sessions, the participant filled in the evaluation questionnaire for the 2D pain drawing. Accordingly, the same questionnaire evaluating the 3D pain drawing was also provided after the end of the last, fourth session. Finally, another participant would be approached on the next scheduled day, and the protocol was repeated.

4 Evaluation Results and Discussion

The analysis of the results consisted of a graphical investigation performed by using statistical software such as Microsoft Excel, and highlighted a general consensus (Fig. 3) that the participants had in respect to the ability of the 3D pain drawing to better visualize and record their pain, as compared to the 2D pain drawing.

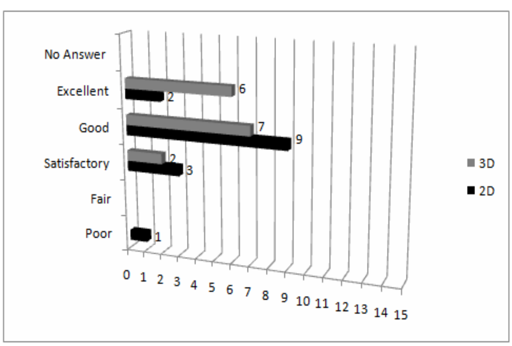


Fig. 3. Overall Performance 2D vs. 3D Pain Drawing

This is generally in line with our expectations. In fact, although opinions about the importance of recording pain information and the usefulness of doing so across time were roughly the same when using both the 2D and 3D drawings, surprisingly, despite the age group of some of the participants, it was found that using the 3D drawing on the laptop was easier than using the paper-based, 2D version (Fig. 4a). Nevertheless, we assume that younger patients will not have these same concerns, a condition that is significantly encouraging with respect to the results.

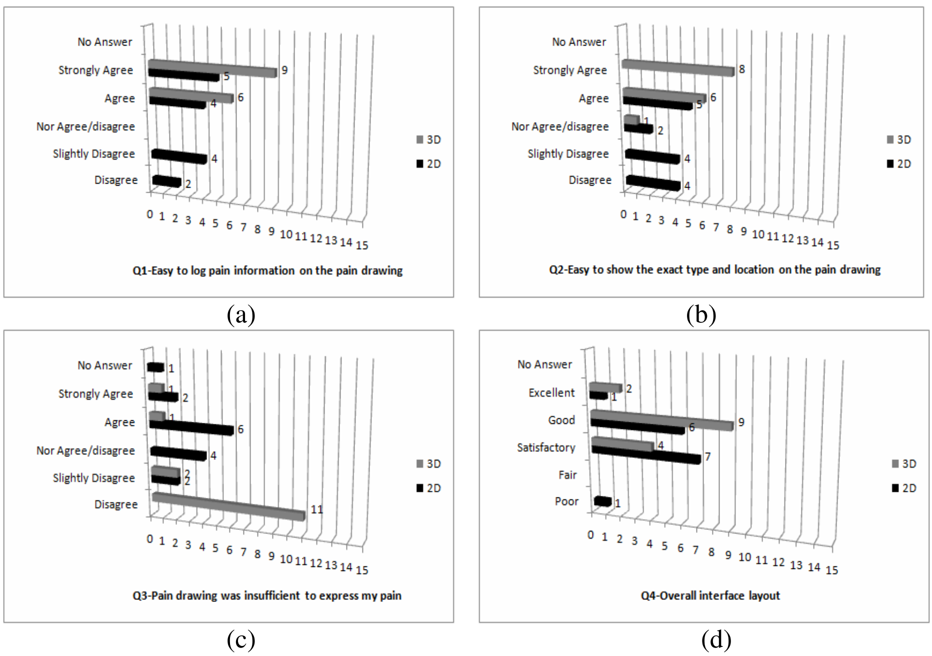


Fig. 4. Bar Graphs of Responses to Evaluation Questions

Specifically, considering that some participants suffered from mobility issues due to the nature of their SCI, it was rational to assume that they would encounter difficulties in using the 3D user interface on the laptop; however, in overall they agreed that the process was a relatively easy one and that the 3D pain drawing was indeed beneficial.

The results with respect to the developed 3D interface's feasibility to visualize pain are also particularly encouraging (Fig. 4b). Considering the limited abilities that the 2D drawing was offering, it comes as no surprise that patients found that showing the *exact* type and location of their pain on the 3D pain drawing was significantly easier than when using the 2D one.

Lastly, our research study has revealed that the general trend was that patients were enthusiastic about the 3D pain drawing, highlighting the wide acceptability and approval of the 3D model's ability to more sufficiently visualize their pain experience

(Fig. 4c). In specific, the majority of the SCI patients that participated in our study appreciated the advantages of the enhanced visualization ability that our 3D model provides by indicating very positive views towards its overall interface layout (Fig. 4d).

5 Conclusions

This study has presented a novel solution for better describing and communicating SCI pain characteristics, which was developed by exploiting the possibilities offered by the advancements in 3D visualization technology. The motivation behind our work lays in the fact that the current pain assessment methods are considered inadequate in their attempt to provide the necessary assistance to the individual with SCI in more accurately describing the pain experience.

As a result, our cohort evaluation with 15 participants, recruited from an Orthopedic Hospital in London, UK, has demonstrated that it is feasible to apply 3D visualization technology in order to more sufficiently assess pain resulting from SCI. In specific, the results of our work have revealed that SCI patients valued the ability of the 3D pain drawing to describe their pain to the medical staff involved in their rehabilitation, by making positive comments, such as *"the 2D drawing was not adequate..",* highlighting in this way the insufficiency of its 2D predecessor.

Moreover, it is anticipated that the use of such technology creates the possibility for patients to become stakeholders in the management of their pain, by allowing first to communicate their pain experience to clinicians in a more perceivable way to the natural environment, and second, to use this visualization ability in order to further understand it. Finally, this study has demonstrated that with the continuous advancements in 3D technology, clinical applications that would utilize such 3D functionality could become an integral part in the rehabilitation of people with some form of disabling pain.

References

1. Broeren, J., Sunnerhagen, K.S., Rydmark, M.: Haptic virtual rehabilitation in stroke: transferring research into clinical practice. Physical Therapy Reviews 14(5), 322–335 (2009)
2. Ciocca, L., Fantini, M., Marchetti, C., Scotti, R., Monaco, C.: Immediate facial rehabilitation in cancer patients using CAD-CAM and rapid prototyping technology: a pilot study. Support Care Cancer 18, 723–728 (2010)
3. Coll, A.M., Ameen, J.R.M., Mead, D.: Postoperative pain assessment tools in day surgery: literature review. Journal of Advanced Nursing 46(2), 124–133 (2004)
4. Felix, E.R., Cruz-Almeida, Y., Widerstrom-Noga, E.G.: Chronic pain after spinal cord injury: What characteristics make some pains more disturbing than others? Journal of Rehabilitation Research and Development 44(5), 703–716 (2007)
5. Felix, E.R., Galoian, K.A., Aarons, C., Brown, M.D., Kearing, S.A., Heiss, U.: Utility of quantitative computerized pain drawings in a sample of spinal stenosis patients. Pain Medicine 11(3), 382–389 (2010)

6. Ghinea, G., Spyridonis, F., Serif, T., Frank, A.O.: 3-D Pain Drawings-Mobile Data Collection Using a PDA. IEEE Transactions on Information Technology in Biomedicine 12(1), 27–33 (2008)
7. Haefeli, M., Elfering, A.: Pain assessment. European Spine Journal 15, 17–24 (2006)
8. Jamison, R.N., Fanciullo, G.J., Baird, J.C.: Usefulness of Pain Drawings in Identifying Real or Imagined Pain: Accuracy of Pain Professionals, Non professionals, and a Decision Model. The Journal of Pain 5(9), 476–482 (2004)
9. Lee, S.J.: Pain measurement: Understanding existing tools and their application in the emergency department. Emergency Medicine 13, 279–287 (2001)
10. Mannion, A.F., Balague, F., Pellise, F., Cedraschi, C.: Pain measurement in patients with low back pain. Nature Clinical Practice Rheumatology 3(11), 610–618 (2007)
11. Mooney, V., Cairns, D., Robertson, J.: A system for evaluating and treating chronic back disability. West J. Med. 124, 370–376 (1976)
12. Ohnmeiss, D.D.: Repeatability of Pain Drawings in a Low Back Pain Population. Spine 25(8), 980–988 (2000)
13. Provenzano, D.A., Fanciullo, G.J., Jamison, R.N., McHugo, G.J., Baird, J.C.: Computer Assessment and Diagnostic Classification of Chronic Pain Patients. Pain Medicine 8 (S3), 167–175 (2007)
14. Ransford, A.O., Cairns, D., Mooney, V.: The Pain Drawing as an Aid to Psychologic Evaluation of Patients with Low-Back pain. Spine 1(2), 127–134 (1976)
15. Samuelsson, K., Larsson, H., Thyberg, M., Tropp, H.: Back Pain and Spinal Deformity-Common Among Wheelchair Users with Spinal Cord Injuries. Scandinavian Journal of Occupational Therapy 3(1), 28–32 (1996)
16. Siddall, P.J., Yezierski, R.P., Loeser, J.D.: Pain Following Spinal Cord Injury: Clinical Features, Prevalence, and Taxonomy. IASP Press, Seattle (2000); IASP Newsletter 3
17. Spinal Injuries Association (SIA): Impact Reports and Financial Statements (2009), `http://www.spinal.co.uk/page/Reports-and-Accounts`
18. Takata, K., Hirotani, H.: Pain drawing in the evaluation of low back pain. International Orthopaedics, 361-366 (1995)
19. Wang, M.Y., O'Shaugnessy, B., Haq, I., Green, B.A.: Pain Following Spinal Cord Injury. Seminars in Neurosurgery 15(1), 99–105 (2004)

Environmental Affordances as a Way to Help in the Design of Videogame Worlds

Elisângela Vilar, Francisco Rebelo, Paulo Noriega, and Luís Teixeira

Ergonomics Laboratory - FMH – Technical University of Lisbon, Estrada da Costa, 1499-002, Cruz Quebrada, Dafundo, Portugal
{elivilar,frebelo,pnoriega,lmteixeira}@fmh.utl.pt

Abstract. Videogame worlds can be read like built environments, so the approaches used to plan real environments may help the design of virtual worlds. In this way, this paper presents a pilot study that the main objective is to investigate affordances of the environment that can influence people's path selection, namely the corridor width. The main hypothesis is that the corridor width will influence people's preference regarding the path they choose in order to escape from a building (in an emergency situation). Stereoscopic images projected in a screen were presented using a constant stimulus method combined with a two-forced choice method to collect user's responses. Findings suggests that there is a tendency to bear right when users are in an "T" intersection where the right and left corridors are equal, and they tend to turn to the larger corridor regardless its direction.

Keywords: Virtual Reality, right/left bias, corridor width, affordances, videogame design.

1 Introduction

The orientation and navigation processes in Virtual Environments (VE) can represent an important aspect for the design of videogame worlds. They also can represent a type of environmental challenge that occurs in many games which require the player to negotiate and remember spatial configurations [1]. All VEs that simulate a complex space (which cannot be viewed from a single vantage point) will have navigation problems, and these problems may be navigators wandering aimlessly when attempting to find a place for the first time or having difficulty relocating places recently visited [2, 3].

Nowadays, many videogames use indoor environments where players travel through corridors, entering rooms and escaping from dangerous situations to accomplish missions. Thus, the design of these virtual buildings has increased importance for the interaction quality and for the users' performance and satisfaction.

However, the more complex a building is, the lower is the developers' knowledge about the people's navigational behavior, since the higher the complexity, the more probable it is that the designer's intuition will fail [4].

If players do not understand where they are and how to get where they need to go, they will get frustrated and instead of moving around the virtual world enjoying the

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 323–331, 2011.
© Springer-Verlag Berlin Heidelberg 2011

game experience, they will look for shortcuts that make more sense for them. Therefore, a game that is not intuitive and easy to use will simply not become popular [5].

Many aspects of videogame worlds (even landscape) can be read as a built environment [6, 7]. So, an approach to study the people's natural movement can be based on the affordances that the environment furnishes or affords the observer [8]. The conscious use of this concept in the design of environments is based on the definition of a set of affordances that express the different priorities and capabilities of a variety of users under various conditions, creating a congruence between what players realize they can do and the activities that they can perform. The formulation of this set of knowledge might be defined studying the people's interaction behavior with indoor environments, mainly with regard to the decisions taken on certain situations.

In this way, this pilot study main objective is to investigate affordances of virtual indoor corridors, namely the corridor width, as well as the right/left bias that may influence people's decision during the interaction with this environment in an emergency situation (escape from a building).

2 Methodology

For this pilot study, in order to investigate the corridor width and the right/left bias which can influence the users' path selection, images of virtual indoor corridors were projected using a stereoscopic projector and 3D glasses. The image sequence was presented using a constant stimulus method combined with a two-forced choice method to collect the users' responses.

The architectural issue considered was the circulation in escape routes, particularly the corridors, which the main affordance is to furnish the passage and the communication among different areas of a building.

2.1 Design of the Experiment

The experiment used a within-subject design and, in order to verify the influence of the independent variable (corridor width) over the people's path selection in an emergency situation, nine different corridors were designed (Fig. 1).

All corridors are composed by two corridors with a "T" intersection type, resulting in two opposite directional choices (left and right).

The left and the right corridors started from 2.00 m of width; the alternative corridors (left/right) had their width increased in 50 cm until reaching 4.00 m of width each.

In this way, the experimental conditions for the independent variable were: one neutral corridor with the same width for left and right corridors (2.00 m) and four corridors varying from 2.50 m to 4.00 m, with increments of 50 cm for both, left and right corridors, alternatively.

Thus, nine experimental conditions were defined. The neutral corridor condition was repeated in order to guarantee the same number of trials for all experimental conditions, resulting in a total of ten stimuli. The stimuli were organized according to the method of constant stimuli and a forced choice between two alternatives method was used to collect answers.

All stimuli were repeated eight times, resulting in eighty trials. These trials were organized in two sequences, a pseudo-random sequence (Sequence 1) and the same sequence in an inverse order (Sequence 2). Thus, two blocks of sequences with forty trials each were presented for all participants, in an alternated order.

The stimulus was a static image of a corridor and had 1400 ms of duration with an inter-stimulus interval varying from 800 to 1000 ms (it was presented to the participants as a gray screen between each corridor image).

The stimulus disappeared as soon as participants push the gamepad button or when the stimulus time was finished. After the end of each stimulus there was still an extra 800 ms time to collet participant's choice.

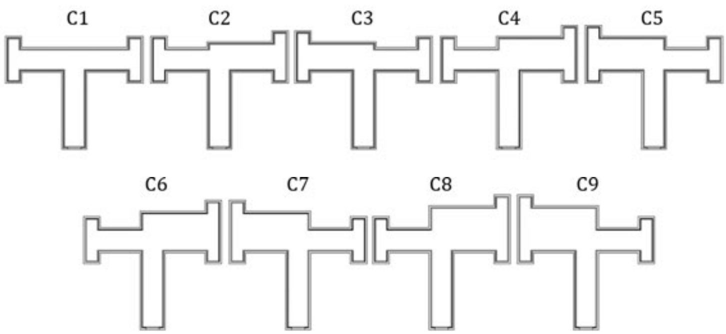


Fig. 1. Nine experimental conditions for corridor width in "T" intersection shape

2.2 Virtual Environments

For the experiment, images of virtual corridors were projected in a screen. Thus, VE's were developed in a way that allows simulating the proposed experimental conditions.

The main requirements to the design of the corridors were:

- Existence of a point for the decision-taking related to the route to follow with two alternative paths (left x right) - Use of "T" intersection types corridors;
- Maintaining the decision point - same distance from the beginning to the decision point for all conditions;
- Constant width for the main corridor and for one of the two alternative paths;
- Variable width in the other alternative path (adding 50cm as width variation);
- Uniform light for all corridor width - Do not use effects of light and shadow;
- Avoid environmental cues - no texturized walls or floor, use of solid colors;
- Insert elements to increase depth perspective – ceiling moldings, wainscoting, baseboards;
- Use of symmetrical elements.

Firstly, 2D plans, which were the base structure of the VE's, were designed using software AutoCad 2008®. These 2D plans were exported to 3D Studio Max® in order to model the 3D environments (9 corridors). In this phase some elements - such as colors on walls and floor, ceiling moldings, wainscoting and baseboards - were inserted in the corridors to increase their realism (Fig. 2). The modeled environments were then exported using plug-in called OgreMax 2.1.2, to be used by the ErgoVR software [9].

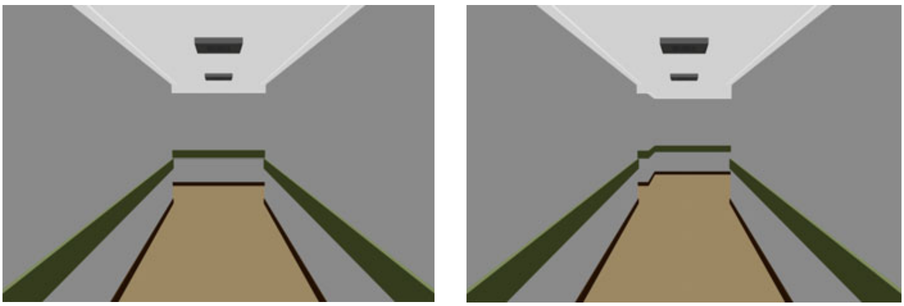


Fig. 2. Neutral corridor condition and 4m width right corridor condition as they were presented by ErgoVR

2.3 Experimental Settings

The Virtual Reality (VR) system used for the experimental tests was comprised by the Lightspeed DepthQ 3D video projector and pair of MacNaughton Inc's APG6000 shutter glasses.

The answers were collected using the Thrustmaster FireStorm Dual Analogue 3 Gamepad, where the functional buttons at right were used to select the answers regarding to the chosen direction (left and right).

The projected images' size was1.72 m (horizontal) x 0.95 m (vertical) and the observation distance (distance between the observer and the projected image) was 1.50m resulting in a 35.2º of vertical Field-of-view (FOV) and 59.6º of horizontal FOV.

All participants remained standing during the experimental session at the same position (marked on the floor) to ensure the same observation distance. The experimenter visually monitored participants and took notes.

2.4 Sample

Eleven subjects, one male and ten females, aged between 20 and 68 years old (mean age = 36; SD= 14.5) participated in this pilot study. Ten participants (91%) declared through a questionnaire the use of the right hand as dominant hand and one (9%) declared the use of the left hand as dominant. All of them were graduated or undergraduated. Half of the participants was firstly assigned to choose the direction in Sequence 1 followed by the sequence 2 and the other half did the inverse order.

The total sample observations comprise of 880 answers that the participants had to choose during Sequence 1 and Sequence 2.

2.5 Procedure

The experimental session was made with each participant separately. Firstly, the participant was asked to sign an Informed Consent Form and advised he/she could stop the experimental session at any time he/she wanted to. The mean duration of each experimental session was 20 minutes. Participants were unaware of the real objective of the experiment. They were told that they should choose one alternative of available paths as fast as possible, as they were in an emergency situation.

The experimental session begun with a brief training session, where the participant had some explanation about the experiment and saw the type of decision he/she should have to take. In this training session participants were told that they must choose between two alternatives of paths, i.e. the one which represents the one they would take if they were escaping from a building in an emergency situation.

The training session also comprised of two practical tests with a sequence of images like those of the experimental test. In the first practical test, the participant was asked to show with his/her hands the both alternatives of corridors he/she was seeing in the image. This was done in order to ensure that the participants realized the alternative of paths that they had in front of them. The second practical test was made in order to make participants familiar with: i) the gamepad buttons which they used to choose their direction and, ii) the time available for their answer since the time used for the training session was the same used for the experimental tests.

Experimental tests began after participants gave their answers in the available time and declared they were comfortable with the gamepad buttons.

For the experimental test, the participant was assigned to the first sequence of 40 trials. Finishing the first sequence and after a 5minutes break, the participant was assigned in to the second sequence.

At the end of the experimental session a demographic questionnaire was applied to collect information such as age, gender, occupation and dominant hand. Participants were also asked to answer questions related to the experimental test.

3 Results and Discussion

The participants' choices regarding the direction they would take for each presented condition considering they were escaping from a building in an emergency situation are presented in Table 1. The results represent the participants' direction choice for all trials (880) and were analyzed globally. There were considered as invalid all missing or impossible answers (when participant pushed a wrong gamepad directional button). Some studies suggest that in ordinary condition (relatively symmetrical buildings) there is a tendency for the public to bear to the right [10,11]. Robinson [10] stated that this trend (about 75% of the visitors bear to the right and about 25% to the left) had been observed in different museums across the United States of America. Scharine and McBeath [11] found that 61.6% of people who had to walk along a corridor formed by a set of paralels bookshelves turned right at the end of the corridor and

38.4% went to the left. Condition C1 represents the neutral condition, with the same width (2.00 m) for all corridors. For the studied situation, results suggest that there is a tendency for the people to prefer the right (64.1%) instead turning to the left (35.9%).

Table 1. Participants' choice for experimental condition regarding the direction

Conditions	Invalid	Left	Right	Total without Invalids	% Invalids	% Left	% Right
C1	6	61	109	170	3,4	35,9	64,1
C2	1	11	76	87	1,1	12,6	87,4
C3	1	60	27	87	1,1	67	31
C4	0	10	78	88	0	11,2	88,6
C5	0	65	23	88	0	74	26,1
C6	0	11	77	88	0	12,5	87,5
C7	1	79	8	87	1,1	90,8	9,2
C8	1	14	73	87	1,1	16,1	83,9
C9	0	69	19	88	0	78,4	21,6

The graphic at Fig. 3 shows the percentage of choices favoring the corridor at the right side. Axis *x* represents the amount of increment added to the width of the corridors. Negative values indicate that these increments were made on the corridor at the left side. The value "0" represents the corridor with no width difference between the corridors on the left or the right sides (neutral condition C1). Axis *y* represents the percentage of choices favoring the corridor on the right side regardless they were the wider or not. Considering the value 0 (C1), we can notice that the 64.1% of the choices were to the corridor on the right side.

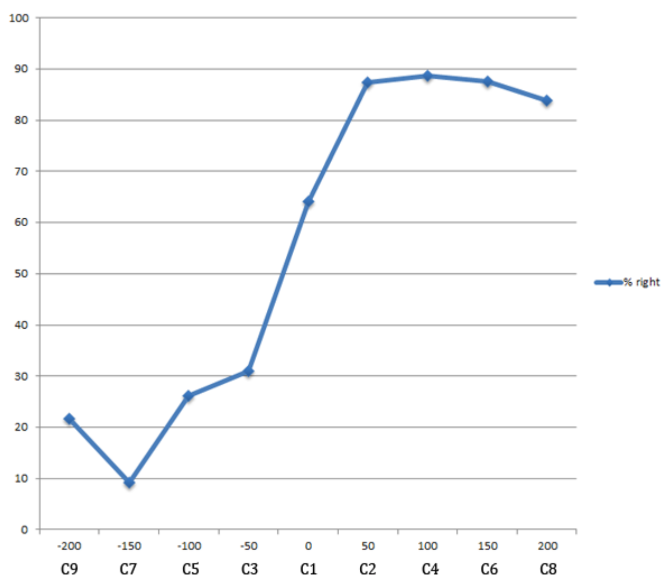


Fig. 3. Graphic with the percentage of choices favoring the corridors at the right side

These findings are in line with previous studies, allowing supporting the hypothesis that people tend to turn right. These findings also contribute to validate the methodology and the setup used in this pilot study to investigate issues considering the variables related to the user's choice regarding the direction.

Fig. 3 also shows that the right bias tends to decrease when the corridor has an increment over its width. The wider the corridor, the lower the tendency to turn right regardless the width. This tendency is contradicted for the largest corridor of the experimental conditions. Curiously, some participants after the experimental session informally reported that there were corridors which were so large that they did not chose to go through them because they will feel like they were alone in that large and empty space. In this way, although the design of this study does not allow verifying this hypothesis, in some way, large empty corridors might provoke bad feelings decreasing their use by visitors. However, despite the plausibility of this hypothesis, the number of participants in this pilot study is small and it may have affected the results.

Considering the corridor width, when the corridor has an increment of 50 cm in width, the effect of this variable is noticed. Condition C3 (Table 1) represents the case where the left corridor had 2.50m of width while the right corridor continued with 2.00m of width. In this case, 67% of participants chose to bear left. The effect of corridor width increases until the increment of 150 cm (C6/C7), as can be seen on Fig. 4.

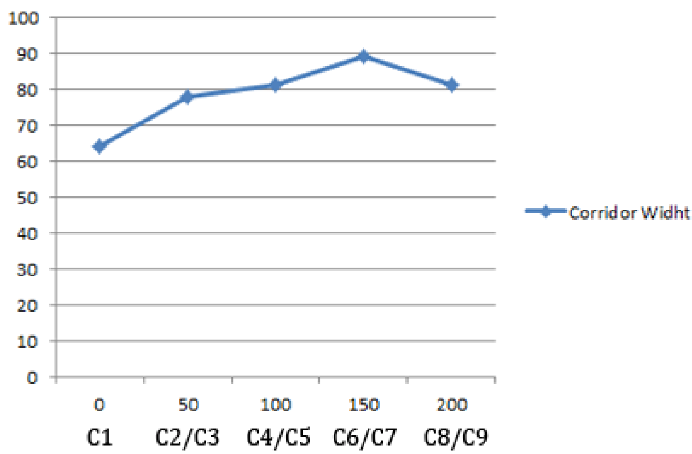


Fig. 4. Graphic with the percentage of choices favoring the large corridors for both sides together

4 Conclusion

A corridor affords passage and communication from one area to another; however when videogame players are confronted with a situation where they have to choose between two options of corridors, some environmental cues can act as attractiveness factor. In this way, the main objective of this pilot study is to verify the hypotheses that people prefer to bear to right and to wider corridors.

Early studies carried out using the real world as interaction environment concluded that in relatively symmetrical buildings or corridor situations, there is a tendency for users to bear to the right [10, 11]. The findings of this pilot study allow corroborating with the early studies results and also indicate that the methodology and the VR setup used in this pilot study are consistent and robust enough to study variables related to the user's choice regarding to the direction.

Regardless left/right bias, when the corridor has an increment of 50cm in the width, the effect of this variable is noticed. Findings of this pilot study allow supporting the hypothesis that users in emergency situation prefer wider corridors. However, data surprisingly show that the choice favoring the wider corridor decreased in the conditions with higher increment value (200cm) suggesting that until certain circumstances the increment in the width of corridor can act as an attraction factor, but if this increment is somewhat higher the result can be the opposite. Despite the plausibility of this hypothesis, the number of participants in this pilot study could be small and it may have affected the results. More research is needed in this area since no studies were found about users directional choices considering variables as width, height, length, lighting, and color.

Affordances are environmental properties which have some meaning to guide the observer's behavior [12, 13], and, according to Venturi and Brown [14], the activities of people into buildings can be seen as patterns of use. In this way, the study of this patterns may be an useful help to game designers to understand people natural movements and predict their behavior, in order to create virtual environments that encompass: a non-obvious way of orientation (opposite to following e.g. arrows, that can make the game boring), an scenario where the user's path selection is under-control and a space that continue being a challenge to the players.

Some questions might be raised with this pilot study. Concerning the sample, this pilot study had a total of 880 observations (80 per participant); however increasing the number of participants and/or the number of trials they performed might produce some effects on the results. In this way, a large sample probably will ensure more robust results in future works.

Another issue to consider is that fixed images projected in a screen, as the setup used in this pilot study, may lack ecological validity. They do not reproduce the natural way of space use; consequently, they do not have high levels of immersion, representing another point that should be improved for future works. When people interacts with an indoor environment they walk and look around to find environmental cues to help with their wayfinding decisions, in this way, using more dynamic scenarios - like videos or animations - with visualization through head-mounted displays and with free visual exploration of the environment will allow a higher level of ecological validity.

Data related to the presence and levels of immersion were not considered for this pilot study, however these variables are also important issues to be investigated. Post-test questionnaires are an easy and useful way to measure presence and level of immersion and their use for future work should be considered.

The contribute of VR systems in the studies related to patterns of space use, namely the people's path selection, can be considered as an important issue since VR systems shown to be an effectiveness way to measure users' decision.

Acknowledgments. This research was supported by the Portuguese Science and Technology Foundation (FCT) grants (PTDC/PSI-PCO/100148/2008 and SFRH/BD/38927/2007).

References

1. McGregor, G.L.: Situations of Play: Patterns of Spatial Use in Videogames. In: Proceedings of DiGRA 2007 International Conference Situated Play, Tokyo, pp. 537–545 (2007), `http://www.digra.org/dl/display_html?chid=07312.05363.pdf` (retrieved)
2. Darken, R.P., Peterson, B.: Spatial Orientation, Wayfinding and Representation. In: Stanney, K.M. (ed.) Handbook of Virtual Environments: design, implementation, and applications, pp. 493–518. Lawrence Erlbaum Associates, New Jersey (2002)
3. Darken, R.P., Sibert, J.L.: Wayfinding strategies and behaviors in large virtual worlds. In: Proceedings of the SIGCHI 1996 conference on Human factors in computing systems, pp. 142–149. ACM, New York (1996)
4. Conroy, R.: Spatial navigation in immersive virtual environments. Unpublished Dissertation, University of London, London (2001)
5. Aarseth, E.: Allegories of Space. In: Borries, F., Walz, S.P., Böttger, M. (eds.) Space Time Play, pp. 44–47. Birkhäuser, Basel (2007)
6. Darken, R.P.: Wayfinding in large-scale virtual worlds. In: ACM SIGCHI 1995, pp. 45–46. ACM, Denver (1995)
7. McGregor, G.L.: Architecture, space and gameplay in World of Warcraft and Battle for Middle Earth 2. In: Proceedings of the CyberGames International conference on Game research and development, Murdoch University, Australia (2006), `http://www.users.on.net/~georgia88/files/Architecture,%20Space%20and%20Gameplay%20-%20Georgia%20Leigh%20McGregor.pdf` (retrieved)
8. Gibson, J.J.: The Ecological Approach to Visual Perception. Lawrence Erlbaum Associates, Boston (1986)
9. Teixeira, L., Rebelo, F., Filgueiras, E.: Human Interaction Data Acquisition Software for Virtual Reality. In: Kaber, D., Boy, G. (eds.) Advances in Cognitive Ergonomics, pp. 793–801. CRC Press, Boca Raton (2010)
10. Robinson, E.S.: The Psychology of Public Educations. American Journal of Public Health 23(2), 123–128 (1933)
11. Scharine, A.A., McBeath, M.K.: Right-Handers and Americans Favor Turning to the Right. Human Factors: The Journal of the Human Factors and Ergonomics Society 44(2), 248–256 (2002)
12. Stoffregen, T.A.: Affordances and Events. Ecological Psychology 12(1), 1–28 (2000)
13. Turvey, M.T.: Affordances and Prospective Control: An Outline of the Ontology. Ecological Psychology 4(3), 173–187 (1992)
14. Venturi, R., Brown, D.S.: Architecture as Signs and Systems: for a mannerist time. Massachusetts. Belknap Press of Harvard University Press, Cambridge (2004)

Analysis of Emergent Use for Wellbeing Service Innovation

Alexandros Yeratziotis[1], Christian Sannemann[2], Johanna Viitanen[3], and Marko Nieminen[3]

[1] Institute for ICT Advancement, Nelson Mandela Metropolitan University, South Africa
[2] Institute of Behavioral Sciences, University of Helsinki, Finland
[3] Software Business and Engineering Institute, Aalto University, Finland
Alexandros.Yeratziotis@nmmu.ac.za,
Christian.Sannemann@helsinki.fi,
{Johanna.Viitanen,Marko.Nieminen}@tkk.fi

Abstract. This paper presents a process of concept development for a new physical activity monitoring device. It forms part of a collaboration project between a Finnish health-technology company and Aalto University's Department of Computer Science and Engineering. There are two main objectives in this project; to develop and validate concepts for the company's product and to evaluate user interfaces that were built on the basis of these concepts. This will result in a set of new ideas for using and improving the service. The paper presents the results from the first and second phases of this three phase project. The first phase focuses on requirements from the health-technology company and how these were considered. Through an iteration process a set of three concepts were derived. In phase 2, these three concepts are visualized via storyboards. In phase 2, prototypes that were designed on the basis of the requirements are also assessed.

Keywords: concept development, obesity and overweight, physical activity monitoring device, user-centered development.

1 Introduction

Physical and mental activities both contribute to the overall well-being and good health of the individual. In the modern era the physical activity of children and adolescents has been degrading alarmingly. This results in serious health concerns at the individual level. For the first time ever, there is reason to believe that an emerging generation has a lower life expectancy than its predecessor [1]. As an example, every third child in the US is obese or overweight. Similarly, the numbers in Europe are also increasing rapidly. In terms of ICT, physical activity monitoring devices and online health services are being developed to assist people in becoming more active. They attempt to promote a healthier lifestyle and motivate people to achieve this. This paper describes an iterative concept development process for a new physical activity monitoring device and online health service. New concepts need to be developed and existing concepts need to be improved. The time span for this project is six months.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 332–341, 2011.
© Springer-Verlag Berlin Heidelberg 2011

1.1 The Problem: Obesity and Overweight

The World Health Organisation [2] defines overweight and obesity as “abnormal or excessive fat accumulation that presents a risk to health”. The most useful population-level measurement to calculate obesity is the Body Mass Index (BMI). The BMI can be applied for both sexes and for all ages of adults. However, one should consider it as a rough estimate because it may not correspond to the same degree of fatness in different individuals. It is understood that a person with a BMI of 30 or higher is considered to be obese, while a person with a BMI equal to or more than 25 is considered overweight.

The probabilities of premature death and disability in adulthood increase with childhood obesity. Being obese or overweight is clearly a considerable risk. It can also contribute to other chronic diseases. These include diabetes, musculoskeletal disorders, cancer and cardiovascular diseases. This is a problem that only used to be associated with high-income countries. However, obesity and overweight are also rising in low- and middle-income countries as well [2]. Obesity and overweight are caused by an energy imbalance. This imbalance relates to calories consumed in comparison to calories expended. The two main factors that contribute to this imbalance are bad diet habits and decreased physical activity [2] [3]. In an attempt to address these two factors, several methods have been suggested to prevent or reduce obesity and overweight [2] [4]. These include achieving an energy balance and healthy weight, increasing physical activity and having a healthy diet. These are methods that can be taken at the individual level.

The problems of obesity and overweight can be tackled with political commitment and collaboration. Stakeholders from the public and private sector have a critical role in the formation of healthy environments. In particular, healthier diet options that are affordable and easily accessible and the promotion of physical activity should be top priorities [2].

1.2 The ICT Approach: Physical Activity Monitoring Devices

As mentioned previously, one solution to reducing or preventing obesity and overweight is to increase physical activity. In an attempt to help people achieve this, stakeholders from the private sector have introduced a variety of physical activity monitoring devices and health services. Most of these devices can be categorized under two core groups; pedometers and accelerometers or multi-sensor activity tracking devices [5].

Pedometers can monitor human activity, like running or walking through the use of motion sensors. They are also referred to as step counters, as they measure the number of continuous steps that a person does. The more recent models have additional capabilities; measuring distance walked, energy expenditure, activity time, heart rate information, GPS (or integrated within cell phones), and BMI information. These devices are portable and wearable and tend to be a good motivator for daily exercise. However, they are still limited in the types of activities that they can monitor. They are mainly used to monitor walking and running activities. In addition, they are well-used in healthcare research and clinical experiments [5]. Some popular devices are the Nike+iPod sport kit, the Nokia 5500 Sports phone, Omron HJ 112, Addidas miCoach Pacer, HJ 720 – ITC and the SportLine 955 pedometer watch. Multi-sensor activity tracking devices overcome some of the limitations found on conventional pedometers and accelerometers. They are also becoming more popular because people require more information regarding their physical activity. As the name implies, these are multiple sensing devices that record physical activities. The sensors are deployed at different areas of the body and are therefore focused on different physiological measurements. As a result, more comprehensive and convincing data of an individual's physical activities can be recorded. Examples of multi-sensor activity tracking devices are the Sensewear armband and the BioTrainer activity monitor [5].

2 Background

Literature review was conducted in the areas of concept development and User-Centered Design (UCD).

2.1 User-Centered Concept Development Process

To integrate users in the formulation of new concepts, the user-centered concept development (UCCD) process was considered. It is defined as a "cross-breed mixture of generic product development, utilizing technological advances and human-centered approach" [6] [7]. The UCCD process has five phases that assist the process of concept development. It helps in formalizing the development into manageable phases for easier adoption. This makes it possible to apply good practices and methods in each phase by taking into account their individual characteristics. The phases and their outcomes are presented in table 1 [8].

Table 1. Phases and outcomes of the UCCD process [8]

User-Centered Concept Development Phases	Outcomes
Phase 1 - Project Commitment: Define user group and con-text, select technology framework, schedule the project	Design brief
Phase 2 - User and Technology Research: Select research methods, conduct user and technology research, analyse data	User tasks and needs description, technology trends and possibilities
Phase 3 - Innovation Sprint: Generate ideas, be creative, do not criticize	Hundreds of ideas
Phase 4 - Iterative Concept Creation and Validation: Select and combine, visualize, validate	Validated concept candidates
Phase 5 - Project Assessment: Evaluate concepts against requirements, collect customer feedback, prioritize concepts and propose future steps	Final concepts, project documentation

2.2 User-Centered Design

The purpose of UCD is to develop products with a high-degree of usability. To achieve this, the user becomes the center of focus in the product development process. Usability is therefore the outcome of applying UCD in the development. UCD is defined as “a user interface design process that focuses on usability goals, user characteristics, environment, tasks, and workflow in the design of an interface; it is an iterative process, where design and evaluation steps are built in from the first stage of projects, through implementation” [9]. To better understand the concept on which UCD is founded, it would be beneficial to first compare it to an alternative approach that is used when developing software products; the system-centered design approach (SCD). The design of a new system in SCD is highly focused on the actual characteristics of the system. For example, designing a product that is to run on a particular platform will evidently influence its design process. This is because the new system will need to be designed in such a manner that it optimizes and fits into the platform for which it is intended [10]. In UCD however, the focus of the design is not based solely on the system characteristics, as is in SCD. Instead, it is based on the fundamental objective to best address users’ needs and their tasks. This is the vehicle that drives the design process. The needs and tasks of users must also be in line with what is stated in the requirements documents. It is even possible to sacrifice certain system efficiency in order to address users’ needs with regards to their interactions with the interface [10].

It is evident that UCD depends on the participation of the intended users of a new software product, throughout the design process. However, the approach that the Finnish health-technology company adopted in their design process tends to adhere more to the SCD process. The project will now attempt to include certain user involvement at the end of design process. This is not the correct manner in which to do UCD but is deemed necessary, even at this point.

2.3 Introduction to New Technology

The device itself is a combination of a multi-sensor activity tracking device and a conventional pedometer and accelerometer. The technology has two key components; an activity logger and a web service. It uses a 3D acceleration transducer and patented algorithms to identify the individual user. The user identification process is based on the actual movements of the individual; each person has their own unique style of movement that distinguishes them from the movements of others. By identifying an individual based on their movements it becomes possible to automatically and accurately determine the type and amount of activity that is being performed by that individual. As with the other devices, this device is also worn by the users during physical activity. Once the users have completed their physical activity they can then connect their device into the USB port of a computer. The data from their last activity will then be uploaded to the web service. Some of the additional capabilities include an activity diary, automatic activity duration recognition, measuring calories burned, training effects (e.g. endurance, strength, etc.), and a long battery life. The most important feature differentiating it from other similar devices is automatic person identification by comparing patterns to samples. In addition, the technology will soon be able to do automatic activity recognition as well, beyond just the standard activities of walking and running (e.g. sleeping, football, etc.).

The problem with the device does not pertain to the functionality of it but instead is focused on the user aspect and usability thereof of the service. This problem is associated with the fact that it is a technology-driven project. This is a controvert case in wellbeing human-computer interaction (HCI) research as in a traditional UCD process, the design is based upon explicit understanding of users, tasks, and environments. However, in this case the potential users will only be involved in later phases of concept development, after the technology and service has already been developed.

In an attempted to improve the usability and user experience of the service, a new project was initiated. It was a special group project for students. The project team consisted of six postgraduate students, some with industry experience, and two supervisors. The team provided a wealth of knowledge, interests and expertise from a variety of research areas, such as usability and user experience for the PC and mobile platforms, security and privacy, usability evaluation methods, software engineering and development, and cognitive sciences.

3 Results

The main objective of the project was to generate concepts for the technology. These included new concepts, as well as improvements on their current services. The project started in October 2010. The planed duration for completion of the project was estimated at six months. Based on these time estimates, the project was divided into three core phases: In Phase 1, concepts would be created by the project team on the basis of stakeholder requirements. In Phase 2, ideas in the form of features and designs would derive from student prototype designs. In Phase 3, concept validation would be conducted. This would assess the concepts and features from Phase 1 and 2 respectively. The final outcome would be a set of master concepts. Phases 1 and 2 will be discussed in this section because it is work that has been completed or is currently under progress. Phase 3 of the project will be mentioned in the Discussion section as it has to be yet conducted.

3.1 Concept Development Activities for This Project

Before the first and second phases of the project are discussed in more detail, an overview of how the UCCD process (from table 1) was applied in this project is presented. This will assist in understanding the objectives of each phase from the process perspective as well. It must also be noted that at present this project is at Phase 4 of the UCCD process. Table 2 presents the outcomes of each phase of the UCCD process with regards to this specific project. The table also displays the status of the phase, which is either “Completed”, “Current work” or “Future work”. The phase of the project that each phase of the UCCD process relates to is also provided.

Table 2. Outcomes of the UCCD process for this project

User-Centered Concept Development Phases	Status	Project Phase
Phase 1: Base set requirements was defined and a project plan was created.	Completed	Phase 1
Phase 2: Literature review in the area of physical activity monitoring devices and services was conducted and research methods were defined. Each member of the project team also conducted research into their own fields that would contribute to the project.	Completed	Phase 1
Phase 3: Based on the research methods in Phase 2, an idea session was conducted that resulted in 115 ideas. The ideas were analysed and combined to develop three concepts.	Completed	Phase 1
Phase 4: The three concepts from Phase 3 need to be visualized. This will be achieved through storyboards.	Current work	Phase 2
Phase 5: A concept validation session will be conducted with users to assess the three concepts from Phase 4. The results will be analysed and the concepts will be improved. A final report will be created for the project.	Future work	Phase 3

3.2 Phase 1: Requirements Elicitation

After several meetings with the company a base set of requirements were determined. These requirements would direct the project scope and goals regarding the concepts. They were particularly significant for the UCCD process. The requirements included: 1. The target audience – the concepts should be focused on the layman adult user. 2. Information presentation – how to display the activity data to the intended users without overloading them with complex data? 3. Useful information – how to make the activity data useful to the intended users? This entails determining what is perceived as useful information by the users. 4. User experience – the concepts should be fun and simple to use for the intended users. 5. Motivation – how to keep intended users motivated and not lose interest in the technology? 6. Integration – how could the technology integrate with other existing services?

A more detailed description of the last part of phase 1, the innovation sprint, will now follow. A day idea session was conducted. For setting the mind before the idea generation process started, members of the team presented interesting findings from their literature reviews as well as discussed the base set requirements. The first method that was applied was the 3-3-5 x 2 technique, a modified version of the 6-3-5 x 2 technique. This modification was due to time constants. The modified technique required three members, thinking up three ideas every five minutes and passing them on for others to develop. As the project team consisted of six members, the ideas were developed by two groups consisting of three members in each. In total 18 initial ideas were constructed all of which were developed twice. Following right after, a brainstorming session was conducted. The ideas from both these techniques were discussed and analyzed and an affinity diagram was developed. After filtering out duplication, the idea session produced a total of 115 raw ideas. The ideas were divided in ten different concept categories within the affinity diagram. These can be regarded as the first-round concept ideas and are displayed in figure 1 as the "10 FIRST-ROUND CONCEPTS". A more in-depth analysis was done after the creation of the ten first-round concept ideas. Some concept ideas were combined and others were improved upon. This resulted in the formulation of eight second-round concept ideas, which are displayed in figure 1 as the "8 SECOND-ROUND CONCEPTS". A selection process on which three concepts would be further analyzed and visualized then followed. This also included discussions with the stakeholders, regarding the concepts. Once again, an in-depth analysis of the eight second-round concepts was conducted. This was done in order to progress to three third-round concepts. The improved concepts that were selected for transformation into visualizations were: 1. Social media game, 2. Instructional and entertaining information, 3. Safety and health device for the elderly. These are displayed in figure 1 as the "3 THIRD-ROUND CONCEPTS".

3.3 Phase 2: Prototyping the Services

The aim of this phase is twofold: to provide useful features and design elements that could be incorporated into the three third-round concepts from Phase 1 and to visualize those concepts for validation in Phase 3. The features and design elements would be derived from student prototypes. These were assignments given to about forty students in an undergraduate course. Based on the concepts and stakeholder requirements from Phase 1, the students were provided with health and well-being scenarios for their prototypes. Once selecting a scenario, students would have to create three types of interfaces for it; CLI (Command Line Interface), GUI (Graphical User Interface) and Web user interface. Although slightly differing from the concepts to be visualized and validated in phase 3, all scenarios were relevant in that they considered three critical requirements: motivation, information presentation and integration with other services. The six scenarios displayed in figure 1 as the "HEALTH & WELL-BEING COURSE SCENARIOS".

The assistant instructor of the course, whom was also a member of the project team, assessed about eighty prototypes to determine the top ten prototypes. The top ten prototypes were then analyzed by the project team. A total of nine interesting ideas resulted from these assessments. Some ideas were: 1. Implementing graphs that display different qualities that are specific to the type of exercise (e.g. boxing is an exercise that is associated with strength, so the user would be interested in information about strength and flexibility). 2. Displaying rewards that would be of interest to the intended users as pictures (e.g. spa days, free newspaper edition). 3. Implementing graphs that compare the user's condition and activity levels (e.g. what is the activity offering the user; endurance or strength, which muscles are being affected by the activity, projections of future progress based on current activity levels, setting activity goals; short- and long-term, displaying achievement ratios of set goals). 4. Displaying the user's physical progress in a picture format (e.g. this could be animations displaying the physical changes to the user's body over a period of time).

The ideas presented above were then visualized so that they can be included into the concept validation process in Phase 3, where possible. This would help determine the effectiveness and usefulness of the design features in each of the three third-round concepts. An illustration summarizing all the activities included in the project's three phases is presented in figure 1. Interconnections between the activities are also provided. At present, the project is in Phase 2: prototyping the service.

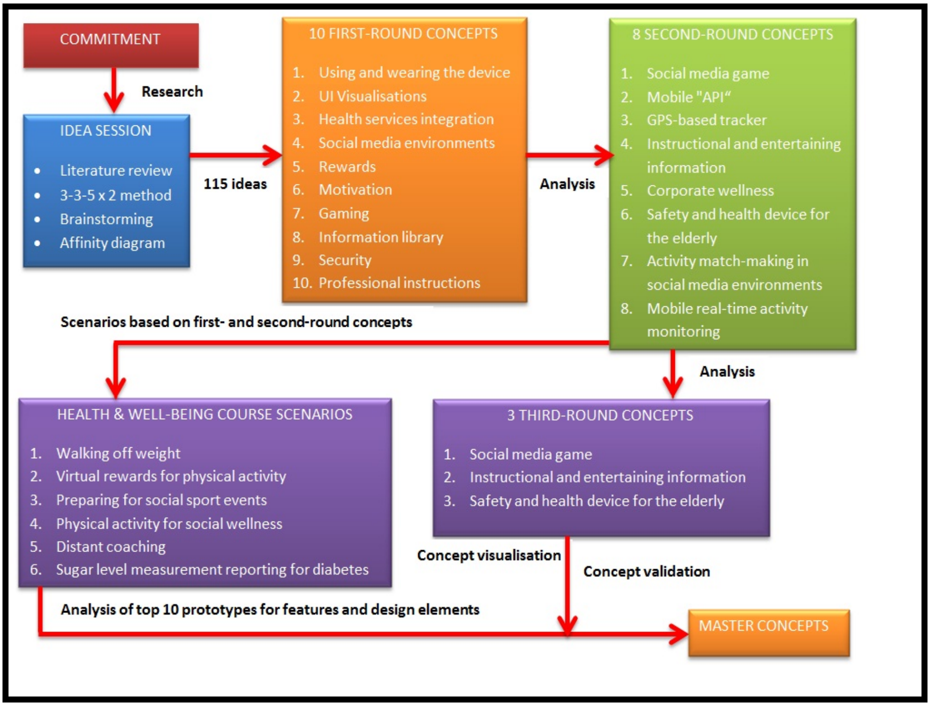


Fig. 1. The three phase process of the project and its activities

4 Discussion

As already stated, in phase 2, storyboard visualizations will be created for the three third-round concepts; Social media game, Instructional and entertaining information, Safety and health device for the elderly. These visualizations will then be used in the concept validation process. A brief description of how the validation will be conducted and the steps involved is provided: 1. Recruit users for testing the visualizations. Three to four users will be recruited for each concept. Therefore, a total of nine to twelve users will be required. The users will use the physical activity monitoring device in order to experience the visualizations in real context settings as well. 2. After testing, a focus group discussion will be conducted for each of the three concepts. The discussion will be focused on three key elements from the storyboards;

motivation, information presentation and integration of services. 3. The data will then be analyzed by the project team. Based on the analysis a final set of master concepts and a report will be provided to the stakeholders.

4.1 Demonstration of Storyboard Visualization for Social Media Game Concept

An initial representation of the storyboard for the Social media game is presented for demonstration purposes. The storyboard is still being improved and visualizations as to how the user is provided with options for physical activity during game play are still under construction. To summarize, the social media game concept is founded on the notion of users that are struggling to progress in a game (figure 2 - left). After a number of failed attempts, users will be provided with the physical activity option (figure 2 - center). If users agree to take the option, they would then temporarily interrupt their game and do some type of physical activity. They would be required to have the physical activity monitoring device with them during their activity sessions (figure 2 - right). Once they have completed their activity, they will then connect their device into the USB port of their computer (figure 3 - left). The activity data is then uploaded to a service, which will provide users with credits based on the amount of exercise done. The user can then use these credits in the game setting to improve their skills (figure 3 - right). This should provide them with the abilities to progress in the game.

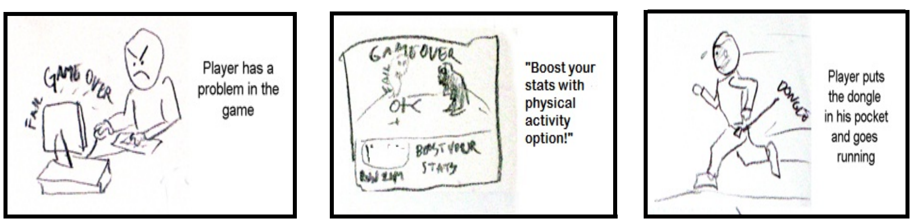


Fig. 2. User is not progressing in the game (left). User is presented with physical activity option (center). User accepts physical activity option and goes for a jog with the device (right).

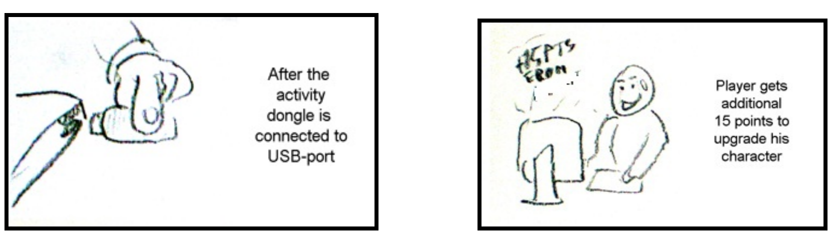


Fig. 3. User returns from jog and connects the device to the computers USB port (left). User receives 15 credits for the jog to improve game character (right).

5 Conclusions

The problem and implications of obesity and overweight are clearly understood. It is also evident that the amounts of people that are facing this problem are constantly rising at an alarming rate. In order to prevent or reduce this problem, several measures have been introduced. One of the measures is for people to increase physical activity in their daily lives. As a result, selections of physical activity monitoring devices are being offered. Their purpose is to help motivate people to exercise more and to provide them with useful activity information. This paper introduces a new physical activity monitoring device.The service for this new physical activity monitoring device was based on a technology-driven project. To improve the usability and user-experience of the service a new project was consequently initiated. The purpose of this project is to improve the current concept services as well as to create new concepts for the technology. o achieve this, it was first necessary to consider the UCD and UCCD processes. Based on the stakeholder requirements, the three main elements that were significant for concept development were motivation, information presentation and integration of services. The project was decided to be conducted in three phases, and is currently in Phase 2: prototyping the services. Phase 3 will initiate once Phase 2 has completed.

Acknowledgments. We thank our colleagues in the special assignment group and the Finnish health-technology company for their collaboration.

References

1. Olshansky, S.J., Passaro, D., Hershow, R.C., Laden, J., Carnes, B.A., Brody, J., Hayflick, L., But, R.N., Allison, D.B., Ludwig, D.S.: A potential decline in life expectancy in the United States in the 21st century. The New England journal of medicine 352(11), 1138 (2005)
2. World Health Organisation: Fact sheet N°311: Obesity and overweight (2006), `http://www.who.int/mediacentre/factsheets/fs311/en/`
3. NHS Choices: Causes of obesity (2010a), `http://www.nhs.uk/Conditions/Obesity/Pages/Causes.aspx`
4. NHS Choices: Treating obesity (2010b), `http://www.nhs.uk/Conditions/Obesity/Pages/Treatment.aspx`
5. Chen, C., Anton, S., Helal, A.: A brief survey of physical activity monitoring devices. University of Florida Tech. Report MPCL-08-09 (2008)
6. International Organization for Standardization: ISO 13407 International Standard: Human-centred design processes for interactive systems, Geneve, Switzerland (1999)
7. Ulrich, K.T., Eppinger, S.D.: Product Design and Development, 3rd edn. Irwin series in marketing. McGraw-Hill, New York (2004)
8. Nieminen, M. P.: Process Methods of User-Centered Concept Development. Licentiate's Thesis. Helsinki University of Technology, Department of Computer Science and Engineering (2006)
9. Henry, S.L.: Just Ask: Integrating Accessibility Throughout Design. ETLawton, Madison, WI (2007) ISBN 978-1430319528
10. Leventhal, L., Barnes, J.: Usability engineering: process, products & examples. Prentice-Hall, Englewood Cliffs (2007)

Visual Innovation through Findings in Perception

Paul Michael Zender

Professor, Director Graduate Studies in Design, University of Cincinnati
School of Design, PO Box 210016, Cincinnati, OH 45221, USA
mike.zender@uc.edu

Abstract. Whether through natural ability or educational training, designers possess advanced knowledge of visual form. Designers acquire most of this special knowledge through experience creating visual objects such as drawings, color symbols, and layouts. Although designers immerse themselves creating effects for visual perception they mostly do so without awareness of the causes of these effects. Designer's form knowledge is more tacit than explicit. Perceptual scientists on the other hand have explicit knowledge of perception of visual form. They have identified physiological processes that directly relate to designers, such as: pop-out, visual illusions, and mental images. The question is less whether designers should be aware of these findings than how the growing understanding of visual perception and cognition can stimulate visual innovation and design practice. This paper reports on initial attempts to integrate this knowledge into design education at the University of Cincinnati.

Keywords: visual perception, visual studies, visual form, tacit knowledge, pop-out, mental imaging.

1 Introduction

There's a saying: you don't know what you don't know. This seems self evidently true. But designer's prove every day that you *can* know what you don't know! For example, many beginning design students are taught by experience that 2 colors can be made to look like 3 (Albers, 1963, p. 78), and many experienced designers are aware that the horizontal and vertical sides of geometric squares don't appear to be the same length (Ruder, 1967, p. 94), yet very few of these designers or students know why these phenomena occur. Designer's can know that form behaves in certain ways in specific circumstances, without knowing how. This may be because design, still emerging from a crafts background has a knowledge base that is more experienced or 'tacit' than propositional or 'explicit.' Tacit knowledge, things we know that we have never put into words and propositions, is valid knowledge. Polanyi illustrates this, saying, "We know a person's face, and can recognize it among a thousand… yet we cannot tell how we recognize a face we know." (Polanyi, 1983, p. 4) Though tacit knowledge may be true knowledge, it is an inadequate base for a discipline (Poggenpohl, 2009, p. 7), because tacit knowledge shrouds concepts in activity and prevents their overt identification, examination, discussion, and verification. For design to advance from craft to discipline some of its tacit

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 342–351, 2011.
© Springer-Verlag Berlin Heidelberg 2011

knowledge must be converted to explicit knowledge. The epistemological basis for this paper is that knowledge in a field grows by converting tacit knowledge into explicit knowledge and that this can happen in design.

2 Visual Perception

We understand much of our world by sight, so much so that when we grasp a thought we often simply say, "I see." Every designer seeing these words relies visual sight and nearly always embodies design solutions in visual form, so much so that a great deal of design education is in fact visual form education. This education commonly happens in studios, with paint and paper, in an experiential learning context where exercises build primarily tacit knowledge. Perceptual scientists on the other hand have recently made explicit many functions of visual perception and cognition that relate directly to designers' manipulation of visual form. Physiological processes of visual perception, all now well-understood by perceptual scientists, are largely unknown by designers. As Nobel Prize winner David Hubel wrote, "it seems unfortunate that people in general and artists in particular should be so insulated from them (findings in visual perception)." (Livingstone, 2002, p. 9) In 2000 perceptual psychologist Colin Ware wrote "Information Visualization: Perception for Design" (Ware, 2000) intending to help bridge the disciplines of perception and design. There he outlined a number of recent findings in visual perception. These advances, empowered by the development of non-invasive brain imaging techniques, define perceptual processes applicable to designers' use of visual form. Three are particularly relevant to the work reported here: pop-out; visual illusions; and mental imaging.

2.1 Pop-Out

Pop-out is a phenomena where some visual forms stand out from surrounding objects very quickly and effortlessly. A pioneer in the study of pop-out, Ann Treisman, described the perceptual mechanism underlying pop-out as 'preattentive processing'. [9] (Treisman, 1985) She suggested that at a very early stage of visual processing, stimulus from our retina are processed rapidly, simultaneously, and 'pre-attentively' so as to identify elemental visual form features, such as: line orientation, curve versus straight line, size, value, color (hue), and blurriness. These primitive visual features are found without the person trying to identify them (hence 'pre-attentive'), and are then processed further in other parts of the brain where they are formed into objects, identified, and acted upon physically or cognitively. The early stage of pre-attentive processing happens very rapidly, in less than 10 msec. (Treisman & Gormican, 1988), and regardless of the number of distracting visual forms, hence 'pop-out.' While the process name 'pre-attentive' has subsequently questioned by Ware, (Ware, 2008) it remains clear that some visual features clearly cause visual forms to immediately stand out from their surroundings, while forms with other visual features take significantly more time and effort to identify. For a designer concerned with hierarchy and communication, the implications of this knowledge are profound.

However, the experience of pop-out is not as simple as identifying pre-attentive features. Many of Treisman's studies isolated just two visual forms and involved a search for one of them. In many of these studies the object with the fewest instances is the one that pops-out. In our own experiments the 'odd-one out' principle suggests that the less frequent object always tends to pop-out from a larger field of related objects. It seems that 'pop-out' is a result of at least two processes: pre-attentively processed early feature detection, and odd or small quantity dominance. These two processes and their interaction are essential to the formation and control visual hierarchy, and hierarchy is a staple of design practice. In his book "Information Architects" Richard Saul Wurman describes hierarchy as one of five ways to organize information. (Wurman, 1997) But hierarchy is even more fundamental than that. It is how designers capture and direct attention. Knowing that hard-wired perceptual processes determine what pops-out first in a visual field equips designers to use forms that will attract the attention of people regardless of age, gender, education, language, or culture.

2.2 Visual Illusions

Optical illusions, art, and their connection to mental processes have been explored widely since the publication of Richard L. Gregory's "Eye and Brain."(Gregory, 1966) More recent work by neurobiologist Margaret Livingstone connects artists' use of value and hue to perceptual process, in particular to the physiological feature center surround.(Livingstone, 2002, p. 122) Citing Stephen Kuffler's discovery in 1953 that retinal ganglion cells are stimulated more by small spots of light than large ones, Livingstone describes how these cells emphasize edges and as a result how a background can change visual objects placed on it. For example, placing a flat gray disk on a rectangular gradient field makes the flat gray disk appear to also be a gray gradient going the opposite direction.

The perceptual experience that results from center surround has been studied for centuries. Michel Chevreul (1786-1889) described how a flat band of color looks different along the edge of a different band of color, particularly if the values are different, one dark and the other light for example (see Figure 5). At about the same time as Kuffler's discovery, Josef Albers' students at Yale University were producing numerous classroom color studies in what was, in effect, an early design research laboratory. Albers reported on these student studies in his groundbreaking treatise "Interaction of Color." (Albers, 1963) There he described the change of the appearance of a color by manipulating the color of its context. Thought Albers probably did not know it, color effects such as "Reverse Grounds Chapter VI – 1 color looks like 2" where a single yellow line is made to look like two very different yellows by modifying the background, or "3 colors appear as 2" where a two single color center squares on two different backgrounds are made to look like their opposite backgrounds, were in fact explorations of center surround perceptual processes. Albers observed, "The true color of the 2 central squares therefore becomes unrecognizable, as it looses its identity." (Albers, 1963, p. 86) Albers was probably unaware of Kuffler's discovery and certainly did not anticipate the body of knowledge that has grown from it. Alber's work affirms that designers can know what, without understanding how, and the lack of further development of Albers' work the past 60 years demonstrates how little progress is made using primarily tacit knowledge. Not knowing how color interactions behave perceptually, designers have failed to build a body of further exploration.

Color illusions are important to designers who use color for any purpose, but particularly for designers who rely on color to encode data for information visualization. When a particular hue on a map is supposed to represent a particular temperature it is a dangerous communication mistake to disregard the background color on which the temperature color will be placed. Edward Tufte has noted the cost of poor information communication in his description of the space shuttle Challenger disaster.(Tufte, 1997, pp. 38-53) This suggests that color illusions, and the perceptual phenomena that cause them, should be well understood by information designers.

2.3 Brain Icons

The existence of mental images, things we see with our minds rather than with our eyes, has been debated. Stephen Kosslyn claims to have settled the debate in his book “Mental Imaging.”(Kosslyn, Thompson, & Ganis, 2006) Kosslyn, professor of psychology and researcher in cognitive psychology and cognitive neuroscience at Harvard University, describes visual perception as a series of subsystems. The earlier of these, the ‘object-properties-processing’ subsystem, is involved with pop-out described above. Objects properties are then passed to the object forming subsystem, then to the associative-memory subsystem where their properties (shape, color, etc.) are compared to the properties of previously known objects. Thus previously seen objects are ‘identified’ and new objects remembered. Key to Kosslyn’s theory is that we can take these remembered images and temporarily push them into the lower subsystems where we can ‘see’ them mentally and perform work on or with them. Key to designers is that Kosslyn believes these memories are stored in an abstracted visual form made up of characteristic visual components. These iconic images have simple distinctive form and typical point of view: a knife has an iconic form: handle and blade, as well as an iconic viewpoint: from the side not from the tip of the blade. These simple forms can have detail added to them. Kosslyn cites a bicycle (p. 144) and notes that this ‘skeletal image’ can have more detail added as needed and if available.

What Kosslyn describes, I believe we could call ‘brain icons:’ simple mental representations of prototypical objects with prototypical viewpoints. Each person, to function in the world, carries a full palette of brain icons corresponding to their experience. When a new object is encountered, or a familiar object with a previously unfamiliar point of view, the person either makes a correct identification or matches the most similar one in memory. This explains popular fascination with photographs of familiar objects from unfamiliar points of view. Shared mental images and points of view are a fundamental process of visual communication. As designed Paul Rand noted, designers “communicate common things in uncommon ways.”

3 Application of Perception in Design Education

These findings are clearly relevant to design. Books by noted perceptual researchers Ware and Livingstone have attempted to make this knowledge available to designers.

(Livingstone, 2002; Ware, 2004, 2008) Perceptual scientist David Hubel notes that this knowledge is not only available, but accessible as well, saying, "I have never had the least doubt that given two hours I could make anyone with a good high-school education fully aware of the main accomplishments of the last half-century of visual science." (Livingstone, 2002, p. 8) Yet, so far, designer and design educators have managed to go on ignorant of how we see what we do. This suggests the question, is knowledge from perceptual science useful for design? For example, can such knowledge make designers better? The premise of this paper is that it can. The question addressed in what follows is how. How can explicit knowledge from the sciences of visual perception and cognition be used to stimulate visual innovation in design education and ultimately in design practice?

3.1 Process

Tacit knowledge, being experienced and assumed and not consciously discussed, is by its nature is hard to identity. Designers aren't aware of what they don't know. But as designers learn about findings in visual perception these can trigger associations with their experience of visual form and thereby raise to conscious thought a formerly implicit understanding. Once the implicit is exposed, it can be transformed to explicit knowledge through rational study. Since 2004 a rational process of exploration has been used at UC with laboratory experiments, some of which have been reported previously (Zender, 2007), and project applications. Several are described below.

3.2 Lab Experiments of Pop-Out

From 2003-2007 approximately 90 experimental studies were designed and conducted by undergraduate students, and from 2008 – 2010 approximately 45 studies were performed by graduate students, all at the University of Cincinnati, and all founded on findings on pop-out from perception science. As reported previously, the method was scientific. (Zender, 2007) Findings in perceptual science were studied. Based on this, questions relevant to design were defined and hypotheses were stated. To test the hypotheses laboratory experiments were designed, conducted, and the results analyzed. The aim was to define basic parameters of preattentively processed / pop-out visual features leading to principles of visual form that could provide a foundation for form-making decisions in design practice.

Ware lists a number of preattentively processed or pop-out features and our experiments have explored nearly all of them, including the hard-to-define feature: blurriness. But over time the most studied features have been color, value, and shape. These have become the focus in part because of the application assigned concurrent to the experiment design (described below under Applications of Visual Illusion).

One recent study by MDes students Gabriel Botkins and Doug Hovecamp addressed the question: does size of object play a role in the ability to perceive gray value, and if so, at what point does it affect perceived gray value and in what way. They hypothesized based on their own tacit knowledge that "as the size of the circles decrease the user will take longer determining the three darkest circles." (unpublished research report) Botkins and Hovecamp designed, and participants took, an interactive test that was taken on a computer and timed in software (Flash). Participants were shown a series of circles arranged in a 5x5 grid of varying values and were asked to identify the 3 darkest circles.

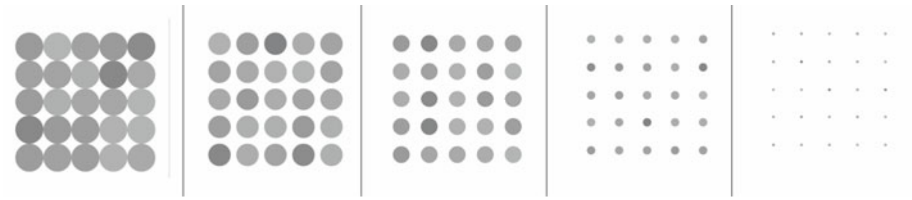


Fig. 1. Representative images from Botekins & Hovekamp. Actual size approx 5 X 5 inches. Circles shown from large to small for illustration purposes, actual test sizes were random.

Each circle's value was determined by software at random within a given range for each set. All circles were uniform in size for each set. Once ten sets were completed to determine the mean time to identify the three darkest values, the circles were reduce in size and the participant asked to complete another set of ten. This continued for 5 sets. The test was conducted on the same computer and monitor to provide a consistent experience at approximately the same distance to maintain the same approximate degree that each circle took up in the participant's field of vision. Thirteen subjects took the test (n=13), nine with corrected vision all of whom were wearing their glasses for the test.

As expected, both the number of incorrect choices and the time to make them increased as size decreased.

Fig. 2. Between the two sizes represented at left, gray value became difficult to judge. Actual size 5 X 5 inches

While the test was definitive, it left unanswered many questions. The large jump in incorrect answers and time to give them between the 45 pixel diameter and the 15 pixel diameter circles invites further study to more precisely define the size at which gray value becomes hard to determine. Though more precision would be helpful, there are more profound questions raised by this study: is the governing factor the absolute size of the circle or the size of the circle as a ratio of the circle and field size; all circles were the same size for each display, would the same result occur if the circles were various sizes; the circles were arranged in a grid, would a random arrangement affect results; and many others.

Questions aside for a moment, if what was found is valid then the uses are significant. Moreover, the finding affirms and enlightens tacit knowledge designers hold that small shapes such as thin rules tend to lose their individual color identity, and that fields of small dots that overlap lose their individual identity altogether and blend with colors of adjacent dots as in printed four-color (CMYK) halftone images and phosphor computer displays (RGB).

Other experiments have revealed that we can see many more variations in light value that dark (see Figure 3).

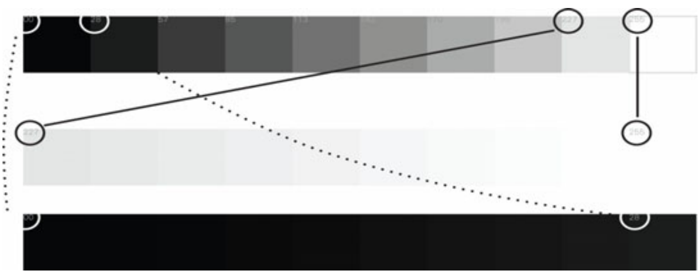


Fig. 3. Mathematically equal gray steps (top row). Ten equally different gray steps inserted between the two lightest steps all are discernable (middle row). Ten equally different gray steps inserted between the two darkest steps none are discernable (bottom row).

We also verified that mathematically equal gray steps do not appear as an even gradation – a finding confirmed by Dona M. Wong in her recent book on information graphics (Wong, 2010, p. 131), that recognizable value differences are smaller when placed on a gray ground, and many others. Such experimental studies are one way of integrating findings in perceptual science into teaching design. But laboratory experiments are just one source for generating explicit design knowledge.

3.3 Project Applications of Color Illusion

At the same time experiments such as that detailed above were being conducted, students in the same class were also given the task of redesigning a color weather map using findings in perceptual science to improve the map's performance. Here the focus was on discovery through applied project design rather than through abstract laboratory experiment. Students were given findings on color perception and the visual illusions: that value is perceptually ordered, anyone can correctly arrange a number of objects of different values; but hue often isn't perceptually ordered: there is no logical sequence to red, green, and blue for example.

Fig. 4. All those asked to arrange gray circles A – B – C in visual sequence (left circles A – C) can easily do so. None of those asked to arrange blue – red – green in visual sequence (right circles A – B – C) can do so.

Students applied this knowledge to design an improved color sequence for their data. Improvement was measured by user test given a simple task such as: where is it the warmest. This different form of study has produced a different kind of results. For example, one observation from this study is that a variety of hues and a strong value sequence support different kinds of user tasks. Tasks surrounding trends and generalizations are best represented by value sequences, whereas tasks finding discrete, specific values are best supported by different hues. The principle could be generalized that hue is good for identifying specific kind: the city at 75 degrees for example; while value is good at showing trend: the coolest region for example.

Various visual illusions were explored as part of these studies. Weather maps typically represent data in bands, with a yellow band perhaps represent 60 degree temperatures and a blue band representing regions at 30 degrees. When sequential data like temperature is more appropriate represented by sequential bands of gray values a new problem occurs: the edge of a gray band looks lighter adjacent to a darker band, and darker adjacent to a lighter band. The bands do not appear flat.

Fig. 5. Flat color bands in map on left exhibit chevruel illusion (example in center). Map on right corrects for this illusion.

The "chevreul illusion," described by Ware in relation to center surround, distorts the impression of the data, emphasizing edges where in fact there is a gradual progression. This stimulated the development of a continual gradient to represent the trend of sequential data like temperature, with interruptions by thin lines to define discrete bands. The value gradient supported 'trend tasks'; the lines supported 'specific-kind' tasks. This is an example where the application of perceptual principles stimulated and guided design innovation.

3.4 Project Application of Brain Icons

In classes related to the design of symbols such as icons, students were given findings on mental imaging and brain icons at the beginning stages of the icon design process. The specific projects in this case have varied. In one instance the project was to create 100+ symbols to help medical students memorize pharmacological interactions – side effects of different drugs; in another case the project was to create a language of symbols to instruct students in resource-poor, low English literacy countries, how to trouble-shoot and repair medical equipment; in another 150+ symbols representing concepts in medicine to facilitate literature review and analysis.

Students were assigned individual objects that were to be part of their visual language of icons. Because some concepts were unfamiliar, the designers researched the object, often including a Google image search. After the research, a game-like activity was undertaken where the object would be named, and participants, sometimes fellow-students and sometimes representatives of the user group, would shout out a necessary feature of their mental image of the object. For example, the object ‘chair’ might elicit ‘seat,’ ‘back,’ ‘legs.’ A probing question ‘how many legs,’ might produce a discourse which resolved into ‘chair = four legs, stool = three legs, office chair = center post + five or more legs.’ Along with the object‘s feature identification, participants were asked for the iconic point-of-view. Again, chair might elicit, ‘side view = yes,’ ‘back view = no,’ ‘top-view = no,’ ‘¾ view from above = yes.’ By this simple means the essential object features and the common points-of-view were identified. Through this process we have seen that some objects have only one iconic point of view, a knife for example. Other objects such as the chair, may have several possible iconic viewpoints. But most all objects have non-canonical viewpoints that make the symbol very difficult if not impossible to recognize: a knife from the tip, a briefcase from the end or bottom, etc.. Lest one think all designers know this well-enough tacitly to avoid serious errors, consider the standard tire inflation pressure symbol required by all USA car manufacturers since 2007. Only 46% of users can identify the symbol as a tire, because it is symbolized in cross-section! Informal surveys suggest that this is not a canonical point-of-view for a tire.

Simply recognizing the tacit knowledge of a brain icon and making it explicit is of great benefit to the student designer. It focuses them on the essentials, freeing them to innovate within the knowledge of the image of object that their users carry in their minds, exploiting that knowledge to their benefit, either to make the object immediately obvious, or if the project calls for it, to make the identity of the object intentionally obscure until the proper moment, like the unfolding of a mystery novel.

4 Conclusion

The lab experiments and applications just described have resulted in some innovative projects. The innovative results were the outcome of this process: investigate findings in visual perception; use these findings to expose personal tacit design knowledge; design experiments and/or exploratory projects to explore the dimensions of the tacit design knowledge; build design principles from experimental findings; deploy the newly developed design principles to design projects; evaluate the projects’ results with users.

Using this process to inspire visual innovation offers several features that make it useful. It is perceptually-based and derived from how people perceive form, not tool-based based on how designers make form. It is process-based and founded on brain function, therefore having the same predictable effect on all people. It is principles-based, not problem or project specific and therefore applicable to a wide variety of situations. It is situation-based and responsive to specific formal tasks at hand, not arbitrarily or randomly creative.

The purpose of this process was to explore means to use findings in perception to stimulate the creation of innovative form solutions. Since the studies were exploratory and not evaluative, we do not know if a 'normal' design process would have produced the same results, or perhaps results that were more innovative. We do not know if knowledge of perception may have impeded the creativity of some students while helping a few others. To answer such questions a control group in an evaluative study is needed.

Clearly, many questions remain, but one thing is clear: the findings of perceptual science are stimulating students at the University of Cincinnati to form innovative solutions to common design problems. This paper proposes that design educators in visual form study findings in perceptual science and then explore how to use these findings to take visual innovation to the next level. It's reasonable to assume that design, part art and part science, can use explicit findings in scientific fields to carefully expose its own tacit knowledge and inform design's investigation of its own designerly phenomena.

References

1. Albers, J.: Interaction of Color, revised and expanded paperback edition. Yale University Press, New Haven (1963)
2. Gregory, R.L.: Eye and Brain: the psychology of seeing, 5th edn. Princeton University Press, Princeton (1966)
3. Kosslyn, S.M., Thompson, W.L., Ganis, G.: The Case for Mental Imagery. Oxford, New York (2006)
4. Livingstone, M.: Vision and Art: The Biology of Seeing. Harry N. Abrams, Inc., New Your, N.Y. (2002)
5. Poggenpohl, S.: Time for Change: Building a Design Discipline. In: Poggenpohl, S., Sato, K. (eds.) Design Integrations Research and Collaboration, pp. 3–22. Intellect, Bristol (2009)
6. Polanyi, M.: The Tacit Dimension. Peter Smith, Gloucester (1983)
7. Ruder, E.: Typography. Arthur Niggli, Ltd, Switzerland (1967)
8. Treisman, A.: Preattentive Processing in Vision. Computer Vision, Graphics, and Image Processing 31, 156–177 (1985)
9. Treisman, A., Gormican, S.: Feature Analysis in early Vision: Evidence From Search Asymmetries. Psychological Review 95(1), 15–48 (1988)
10. Tufte, E.R.: Visual Explanations Images and Quantities, Evidence and Narrative. Graphics Press, Cheshire (1997)
11. Ware, C.: Information Visualization: Perception for Design, 1st edn. Morgan Kaufman, San Francisco (2000)
12. Ware, C.: Information Visualization: Perception for Design, 2nd edn. Morgan Kaufman, San Francisco (2004)
13. Ware, C.: Visual Thinking For Design. Morgan Kaufman, New York (2008)
14. Wong, D.M.: The Wall Street Journal Guide to Information Graphics: the dos and donts of presenting data, facts, and figures. W. W. Norton & Company, New York (2010)
15. Wurman, R.S.: Information Architects. Graphis Inc., New York (1997)
16. Zender, M.: Basic Design Research: Parameters Of Visual Form. Paper presented at the International Association of Societies of Design research IASDR 2007, Hong Kong, CN (2007)

To Develop Viable Human Factors Engineering Methods for Improved Industrial Use

Jonas Andersson[1], Lars-Ola Bligård[1], Anna-Lisa Osvalder[1],
Mikko J. Rissanen[2], and Sanjay Tripathi[3]

[1] Chalmers University of Technology
412 96 Göteborg, Sweden
[2,3] Industrial Software Systems, ABB Corporate Research
[2] Forskargränd 7, 721 78 Västerås, Sweden
[3] Whitefield Road, Bangalore-43, India
{jonas.andersson,lars-ola.bligard,
anna-lisa.osvalder}@chalmers.se, mikko.rissanen@se.abb.com,
sanjay.tripathi@in.abb.com

Abstract. Human factors engineering methodology is important for design of complex systems, such as control rooms and distributed control systems. Available methodologies are however seldom adapted to industrial needs, which limits the use of the existing human factors engineering research base. In this paper we argue that human factors engineering methods have to be developed and adapted to the engineer working under industrial project constraints. Otherwise human factors engineering is unlikely to achieve a broad industrial impact. The paper suggests how the industrial viability of methods can be improved by applying a use centered approach to method development.

Keywords: human factors engineering methodology, method development, complex systems.

1 Introduction

The field of human factors engineering (HFE) is defined as the application of knowledge about human behaviour, limitations and other characteristics to the design of for example tools and systems to achieve effective human use [1]. Within this field, research efforts have for decades been directed towards the design of complex socio-technical systems to promote safety, efficiency and productivity in human work. An example of where the application of HFE is important is control room operations in various domains, such as energy and process industry.

A subset of HFE is cognitive systems engineering (CSE) [2,3]. The methodology within CSE has the potency to handle complexity. The methodology is therefore applicable in development of control room related technology, e.g. development of distributed control systems and in particular in the design of graphical user interfaces. The research efforts within CSE have provided frameworks and methods (e.g. cognitive work analysis (CWA) [4]) to arrive at solutions for improved human supervisory control. The examples of how cognitive work analysis can be used are

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 355–362, 2011.
© Springer-Verlag Berlin Heidelberg 2011

growing [5] but the industrial adoption is still scarce. A reason for this is that while CSE as such aims for useful, usable and understandable systems, the methods that are developed within CSE still need improvement to reach a wide industrial usage. As long as CSE methods are difficult to adapt to industrial project constraints, a broad industrial impact is unlikely to occur.

We argue that a large part of the problem seems to lie in how HFE methods in general are developed by researchers for research purposes, but that many of the methods are unsuccessful in meeting the needs in industrial practice. For methods used within CSE this problem become significant due to the reliance on the methodology to address complex problems in complex work domains. To solve this problem collaboration between academia and industry is necessary to bridge theory and practice. The *user* is always considered in HFE in product development. However, the users of HFE methods, e.g. engineers working in industry are given too little consideration in method development.

The paper elaborates on the importance of viable human factors engineering methods for utilization outside the world of academia. The problems are explained using activity theory and by using a product development approach a suggestion for solution is provided. The purpose of this paper is to outline how human factors engineering methods can be adapted or developed to fit to industrial project constraints and to discuss how the responsibility of method development should be shared between academia and industry. A special focus is given to methodologies that are useful in the engineering of complex systems. In the paper, cognitive work analysis is used as an example.

2 The Limited Use of HFE-Methods in Industry

In this section the problem of methods not being adapted to industrial needs is explained by using activity theory. Further the usability and utility aspects of HFE methods are described and the challenges for viable methods in industry are outlined.

2.1 Methods as Artifacts

The question of why methods are not developed considering engineering practicality can be clarified using an activity theoretical model. Activity theory provides a way to describe the relation between the user, here called “operator” for clarity, and the artifact [6,7]. A subset of the theory is the triangular relation between artifact (means), task (ends) and operator (subject), where the operator uses the artifact to achieve his/her ends (top triangle in figure 1).

Engelbrektsson [8] argued that there is a similar triangular relation between the engineer, the artifact and the method. The engineer uses the method to achieve his/her ends, i.e. develop an artifact. Engelbrektsson further stated that the engineer considers the artifact to be an end itself, rather than seeing the artifact as a mean that the operator needs to achieve his/her relative end. This is a reason for artifacts being developed with poor adaptation to the operator. This can be illustrated using two interconnected triangles (the two top triangles in figure 1). Engelbrektsson argues that part of the solution to poor operator adaptation lies in the direct contact between the engineer and the operator as he/she uses the artifact. This is essential to acquire the understanding of how the operator uses the artifact to reach his or her goals. A way to accomplish this is to apply a use centered approach.

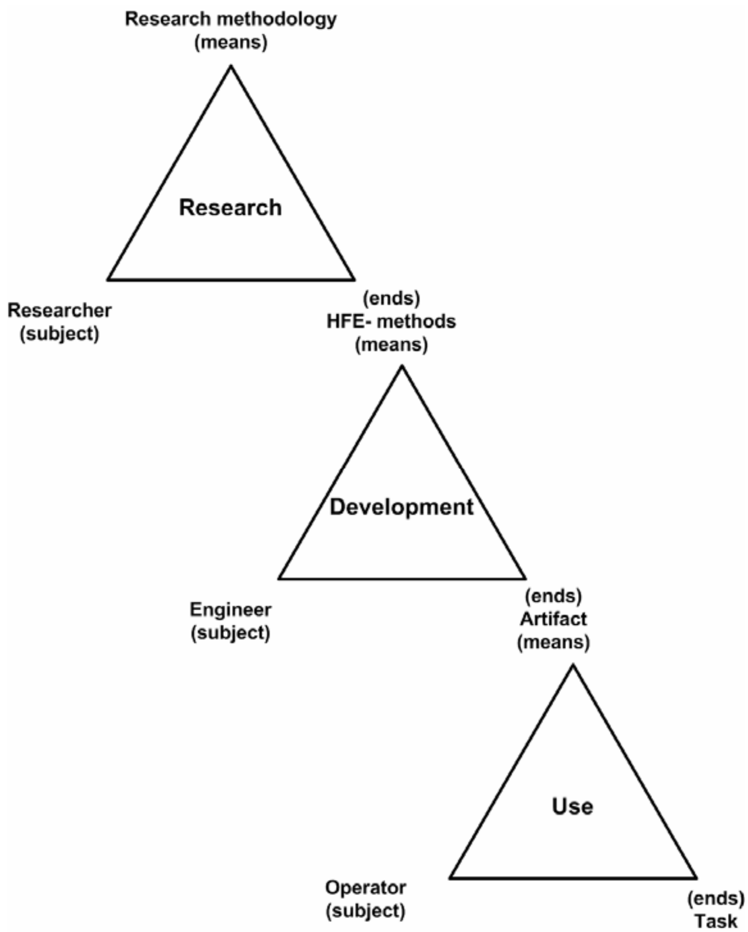


Fig. 1. The triangular relation between research, industrial development and use

According to professor MariAnne Karlsson (Division of Design & Human Factors, Chalmers University of Technology, pers. comm.), the HFE researcher can also be included in this model. The HFE-researcher uses research methodology (means) to develop HFE methods. The problem is that the methods are seldom adapted to industrial practice. Researchers often see their method as an end for research purposes, rather than as a mean for an engineer working under industrial project constraints. All this argumentation can be illustrated by combining three adjacent triangles (figure 1). Our conclusion is that when researchers develop methods, it is important to acknowledge the engineer in industry as the user of the methods if a wider industrial impact of HFE methods should be achieved. The researcher that develops methods must therefore have direct contact with the engineer as he or she uses the method. This is essential to acquire the understanding of how the engineer uses the method to develop new artifacts. Researchers need to apply the same use centered approach when developing methods as the engineers do to achieve useful artifacts. In conclusion, methodological research and development should have a human factors perspective on the method development itself, where the engineer is seen as the end user. Otherwise a wide industrial adoption of HFE is unlikely to occur.

2.2 Usability and Utility of HFE-methods

If an artifact is to be used and provide a positive user experience, let be a product or a method, it has to be useful. Considering usefulness, two characteristics are of importance; utility and usability [9]. In a method development context, a method with good utility would mean that it could provide the results that are needed for further product development. A method with good usability means that it is practical and doable within the company's internal working procedures and fits well within project constraints.

We argue that the focus of present research efforts in method development is mainly directed towards the *utility* of methods. This is not unexpected since utility in terms of getting results from the method would naturally be the first aspect to consider. However, since methods are mainly developed in the research community the *usability* of methods for use in an industrial context is seldom considered. This indicates the need for a use centered approach in method development.

2.3 Challenges For Viable Methods in Industry

Industry is requesting methods that are feasible to implement and at the same time are able to handle complex problems [10]. The question that emerges is how such methods are to be developed? To reach the goal of viable methods the responsibility has to be shared between academia as method developers and industry as users of these methods.

There are several challenges that can be expected in development of methods for industrial use. The methods used in development of, for example, graphical user interfaces for distributed control systems must be able to account for the inherent complexity of the sociotechnical system and the customizability of each installation. It is a challenge to develop easy to use methods for design of advanced systems. Designing simplistic methods for complex problems is a contradiction that has to be balanced.

Another problem is that the human factors engineering knowledge in industry often is lacking in order for advanced analysis work to be conducted. If industry had people with sufficient human factors engineering knowledge, then methods could also be made more sophisticated. In other engineering disciplines it is not expected that tools should be mastered without expertise (e.g. a proficient mechanical strength analysis cannot be expected just because you have a computerized tool for finite element analysis at hand). Human factors engineering expertise is needed and industry has to acknowledge HFE as an *engineering* discipline among other disciplines, such as mechanical and electrical engineering.

In the cases where human factors engineering methods are used in industry, there are still challenges with how to integrate a HFE methodology into the systems engineering process. The output of a method has to be useful and provide relevant information that can for example be transferred from a human factors engineer to a programmer. This requires good integration between the human factors engineering discipline and other engineering disciplines in the product development process. Hence, the understanding of how the methods are used in an industrial context is essential in the method development.

3 Obstacles for Applying Complex HFE in Industry – The CWA Example

There is a range of existing frameworks and methods within CSE [11]. One of the most elaborated and methodologically structured frameworks is CWA. CWA has the potential to cope with complex problems that companies such as developers of distributed control systems are facing. Examples have been given of the practical applications of CWA [5], but the examples provided are mainly performed in a research rather than business related context. This justifies the question if CWA is possible to apply in the product development processes of a company acting under business constraints. Vicente [4] acknowledges that in order for CWA to be used in industry the great effort to conduct a CWA must be overcome. He also states that "*to conduct a CWA for a complex sociotechnical system could take an incredible amount of effort, more effort than many companies are willing to invest*" [4]. To aid CWA Vicente suggests two ways to facilitate the accomplishment of CWA. The first suggestion is the development of computer-based tools to support the CWA work. Jenkins et al [12] developed such a tool in the "CWA Tool". This initiative is promising, however the tool is still in the development stage and not a commercial product. The second suggestion is the development of "prototypical template analyses" for different application domains [4]. The idea of such templates is to capture the constraints that are common to a certain domain, but allow the analyst to provide the details, thereby saving time in the analysis. The authors have however not found any publications that reports of the development of such templates (although they might be used unreported of in applied projects).

Another problem is that although the concepts of CWA have been thoroughly described (e.g. [2,4]), practical descriptions of the methodological steps for full-scale industrial projects are still limited. Such descriptions must allow for relatively novice users of CWA to conduct an analysis if the framework is to gain a wide industrial impact. The methodology can however not be simplified to the extent that it can no longer account for the complexity of the work domain. This poses a balancing challenge for further method development.

As a response to this balancing act, we propose an approach to improve the prospect of a wider use of methodologies such as for example CWA in industry. In order for a method to be adopted in industry, the method has to be feasible in terms of available resources. For example, how can a methodology such as CWA be scaled to fit available resources? Also guidance on how to adapt the resolution of the CWA steps is needed.

In general, methods have to provide useful results and it has to be well integrated in the company's own product development process. In other words, it is essential to find a way to adapt the methods to the needs of the engineers'. To accomplish this the method can be seen as an artifact that can be adapted by applying a product development process.

4 An Approach to Adapt HFE Methodology to Industrial Needs

To reach viable methods useful in industrial projects the usability of HFE methods has to be improved by applying a use centered approach. This can be achieved by focusing on the user needs in method development. By viewing the method as an artifact that can be run through a product development process the practical and methodological needs of the engineers working in industry can be taken into account. Here we outline what such a method development process might look like.

The process can be simplified to three basic steps; identification of needs, development and implementation [13]. The first step of identifying needs should be to define what kinds of methods are needed, e.g. should it be used for planning, data collection, analysis, design or evaluation or a combination of these? This is important so that the scope of the method is clear from the beginning. Further, the company's product development process has to be described. This is made in order to map out the external constraints that will affect how the method will be used. By collecting the needs of the engineers as users of methods this can be made in practice. The factors that influence the use of methods are for example the education and experience of the engineers and the laws and regulations within a specific domain. Also organizational preconditions (e.g. resources) in the work context should be taken into consideration. In this aspect, the method's ability to reflect the domain (i.e. the engineering problem at hand) will determine the method's utility. The method's ability to fit into a company specific work procedures will however determine its' usability.

The next step is to develop a new method or adapt an existing method to meet the identified needs. This is probably the most difficult part where simplicity of use has to be balanced against methodological thoroughness. In this aspect the simplicity relates to the usability while thoroughness relates to the utility of the method. In this step existing methods should also be evaluated to see if they meet the identified needs. Methods need to be made scalable to fit different project constraints in terms of time frames and resources. This also stresses the importance of how to show what is being sacrificed if a smaller methodological effort is chosen, i.e. how the resolution of the method can be adjusted and how the effect of a change in resolution can be indicated. Another aspect is the adaptability of methods to different preconditions, i.e. work domain, type of project and questions to be investigated with the method. Adaptability can either be achieved through development of individual methods that can be adapted to various industrial preconditions, or through creation of methodological packages with a structured guidance in what methods to choose depending on the project preconditions and scope of the project. The needs can be different within a company and will also differ between companies since the demands for methods vary, e.g. depending on what the customers require.

Thirdly, the method is to be implemented which implies that the method has to fit with the existing work procedures in the company. This is important to make the method usable in daily engineering work. In this step the documentation and description of procedures becomes central together with courses for industrial practitioners, pedagogical exercises and computer-based aids.

An example of an existing methodology that has been developed with special focus on fitting into an industrial systems engineering context is Applied Cognitive Work Analysis (ACWA) [14,15]. Although the ACWA methodology is well described, how to achieve methodological practicality and adaptation to an existing systems engineering process within a company in general is not in focus.

5 Discussion

The expected effect of developing methods by using a product development process is that methods will work well in a specific industrial context. This will result in specialized methods that are adapted to a specific industrial setting and to a specific company. This means that a wide area of application is traded for functionality, i.e. that the method actually works where it is needed. The result is a greater variety of more specialized methods.

The intention of this approach is primarily to make existing methods available for industrial use in the same way that existing technology is made available through product development. We believe that this approach can be applied in a straightforward manner since there is much available knowledge regarding product development. The product development process may however need adaptation to fit the development of methods instead of development of products.

We do not suggest that using a product development process should be used in all method development. But if the goal is to reach a wider industrial usage this approach can be a step in the right direction. When results from methodological research should be used in industry, the engineers as users of methods have to be considered.

There are also important differences between research methods and engineering methods. In this paper engineering methods have been addressed. While research methods in general try to come as close as possible to an objective truth, engineering methods can settle with being “good enough” since precision is not required to the same extent. Also, the time available in industrial projects is always limited which means that methods have to be time efficient.

It should also be acknowledged that if human factors engineering is to reach a wider impact in industry, there are more issues to be addressed than just creating viable methods. Firstly, the cost benefit of human factors efforts has to be recognized in order to motivate investment of resources. This might need organizational change and perhaps also a change of mindset to avoid short-term thinking.

To promote HFE in industrial product development in the long run, we also believe that HFE methods have to bring a positive user experience for the engineer performing the analysis. Positive in this sense lies in terms of smoothly arriving at relevant results and reaching project goals. By providing a viable method the engineer will have a positive methodological experience that will inspire both learning and further use of human factors engineering methods.

6 Conclusion

To achieve a fruitful collaboration in development of viable human factors engineering methods, academia and industry have to meet half way. Industry cannot expect methods adapted to their needs without engagement and will to improve their knowledge. Neither can academia expect a wider impact of HFE if methods are too difficult to use. Primarily, academia should put more effort into creating usable methods when HFE methods are developed and improved. Industry on the other hand, should recognize that human factors engineering is an engineering discipline that requires its own in-house specialists.

References

1. Chapanis, A.: Some reflections on progress. In: Human Factors Society 20th Meeting, Santa Monica, CA, pp. 1–8 (1985)
2. Rasmussen, J., Mark Pejtersen, A., Goodstein, L.P.: Cognitive systems engineering. Wiley, New York (1994)
3. Hollnagel, E., Woods, D.D.: Joint cognitive systems: foundations of cognitive systems engineering. CRC Press, Boca Raton (2005)
4. Vicente, K.J.: Cognitive work analysis: toward safe, productive, and healthy computer-based work. Lawrence Erlbaum Associates, Mahwah (1999)
5. Bisantz, A.M., Burns, C.M.: Applications of cognitive work analysis. CRC Press, Boca Raton (2009)
6. Vygotskij, L.S., Kozulin, A.: Thought and language. MIT Press, Cambridge (1986)
7. Kuuti, K.: Activity Theory as a Potetial Framework for Human-Computer Interaction Research. In: Nardi, B.A. (ed.) Context and consciousness: activity theory and human-computer interaction, pp. 17–44. MIT Press, Cambridge (1996)
8. Engelbrektsson, P.: Enabling the user: exploring methodological effects on user requirements elicitation. Chalmers University of Technology, Göteborg (2004)
9. Nielsen, J.: Usability engineering. AP Professional, Boston (1993)
10. Breen, S., Enkerud, T., Husoy, K.: Applying usability engineering in ABB. In: Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles, pp. 491–492. ACM, New York (2006)
11. Militello, L.G., Dominguez, C.O., Lintern, G., Klein, G.: The role of cognitive systems engineering in the systems engineering design process. Systems Engineering 13(3), 261–273 (2010)
12. Jenkins, D.P., Stanton, N.A., Salmon, P.M., Walker, G.H.: Cognitive work analysis: coping with complexity. Ashgate, Surrey (2008)
13. Ulrich, K.T., Eppinger, S.D.: Product design and development. McGraw-Hill, New York (2004)
14. Gualtieri, J.W., Szymczak, S., Elm, W.C.: Cognitive System Engineering - Based Design: Alchemy or Engineering. In: Human Factors and Ergonomics Society Annual Meeting Proceedings, pp. 254–258 (2005)
15. Elm, W.C., Potter, S.S., Gualtieri, J.W., Easter, J.R., Roth, E.M.: Applied Cognitive Work Analysis: A Pragmatic Methodology for Designing Revolutionary Cognitive Affordances. In: Hollnagel, E. (ed.) Handbook of cognitive task design, pp. 357–382. Lawrence Erlbaum Publishers, Mahwah (2003)

What Sustainability Brings to the Research Table

Niyati Bedekar and Sally Lawler Kennedy

SAP Labs, 3410 Hillview Avenue, Palo Alto, CA 94304 USA
{Niyati.Bedekar,Sally.Lawler.Kennedy}@sap.com

Abstract. Sustainability is becoming increasingly important in today's corporate world. Businesses with dedicated sustainability programs are seeing a positive impact on their bottom-line, employee satisfaction as well as their brand image. However, sustainability is ever changing and businesses need to keep up with these changes for future success of their organizations. The evolving nature of this field offers great prospects for user researchers; particularly for those working in an enterprise software company such as SAP that offers a suite of sustainability products to help companies run their businesses more sustainably. Drawing examples from our research this paper will discuss the new roles, responsibilities, and unique opportunities user researchers have; and how we can bring about a positive impact in the field of sustainability.

Keywords: sustainability, user research, opportunities.

1 Introduction

Sustainable development was first defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs." [1] While organizations may have earlier focused on addressing environmental impacts alone, today it is vital for them to consider the social and economic aspects of sustainability as well. According to Giselle Weybrect, "...we are quickly moving away from a '*sustainability is all about saving the world but not about business'* mentality into the *'using sustainability to strengthen my business while also having a positive impact on society'* one." [2] Many global corporate executives are including sustainability objectives in their top priorities. [3] The following figure shows the survey (conducted in 2007) results indicating the priorities of these executives for the next 5 years.

There is great pressure on companies to divulge information on how they are addressing social concerns such as their fair trade practices etc., economic and environmental aspects of sustainability in their business operations. The demands are coming from shareholders, employees, consumers, customers, governments, as well as the society at large.

To address these needs, SAP has a suite of sustainability products to help gain insight into operations and enable the creation of viable business strategies for sustainability. The products range from Carbon Impact, Environmental Health and Safety, Product Safety and Product Compliance to Sustainability Performance Management and People Health and Safety etc. The User Experience (UX) team is involved in conducting user research and creating designs for these products.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 363–367, 2011.
© Springer-Verlag Berlin Heidelberg 2011

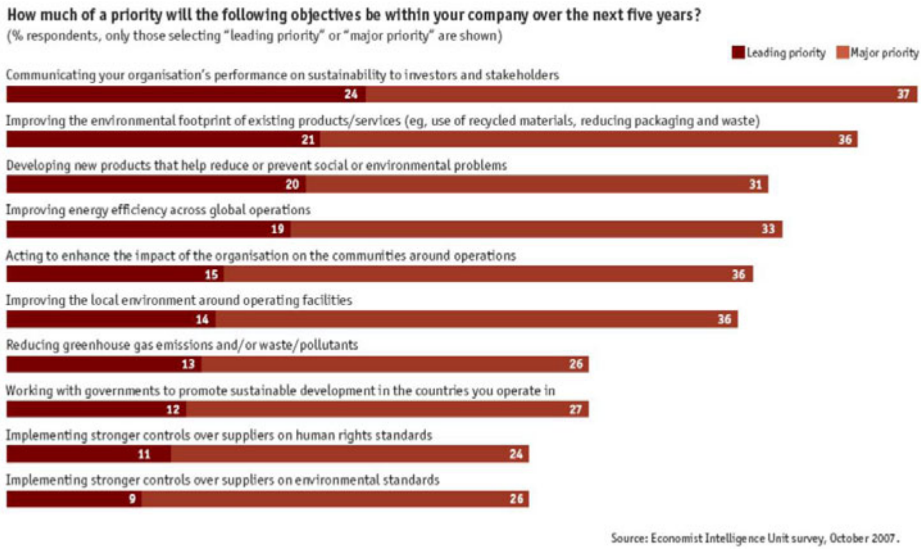


Fig. 1. Responses to one of questions in the Economist Intelligence Unit Survey, October 2007

The SAP UX team is located in multiple locations across the world. The ideas and examples used in this paper were gathered from our multi-country user research focused on sustainability conducted from June 2009 to December 2010. The user research activities included surveys, interviews, field research, contextual inquiries, user interface walkthroughs, formative usability studies etc. The user researchers were based out of Palo Alto, CA, Walldorf, Germany and Sau Paulo, Brazil while the product teams were globally distributed.

In the following two sections we will discuss how user researchers are uniquely poised with opportunities and added responsibilities to help shape the evolving field of sustainability.

2 Roles and Opportunities

2.1 Sustainability Is Beyond Compliance

Many companies have historically looked at corporate sustainability as having to do only with meeting regulatory compliance requirements. We can help companies understand that *sustainability is not just about compliance,* it is a whole new way to think and run your business. In April 2010 Reuters' Environment Forum identified that the number one trend in sustainable business was: "*A deeper understanding of what sustainability means*". Companies are asking the question "What does it mean to be sustainable?" when defining their corporate sustainability strategy. For many companies there was the belief that "sustainable development" will increase their operational costs. However is has been shown that corporate sustainability is a key driver of innovation and can actually positively impact the company's bottom-line in the long run by lowering their costs and increasing their revenues. [4] This holds true for companies like GE, DuPont Wal-Mart that have adopted aggressive sustainability strategies with positive financial gains.

In meeting with different customers during our research, we realized that they were at different stages in their implementation of sustainability programs and practices; many of them strictly focused on meeting regulatory compliance needs while others had started thinking of ways to make sustainability pervasive throughout their organizations. By using a Sustainability Maturity Model such as shown in Figure 2 user researchers can assess the company's sustainability status. Identifying where a company is at enables the user researcher to then collaborate and brainstorm on how to advance a company to be more sustainable.

During interviews with safety managers at one of our multi-national mining customers, it was apparent that their immediate need was to be compliant with MSHA[1] regulations for their US mines. In our discussions, we learned that they desired to be more proactive and move beyond compliance; however, they did not have the tools or technology to make that happen. By understanding what stage the company is currently at, we can help them think beyond the limitations of existing technologies and create a radical vision of how to run their business in order to become a more sustainably mature company.

Level 1	Level 2	Level 3	Level 4	Level 5
Inactive Non-Compliance	**Basic Compliance**	**Beyond Compliance**	**Integrated Sustainability**	**Sustainability Leadership**
- Managers don't know or don't care about environmental and social issues - Regulations and standards do not apply or are ignored - No sustainability efforts	- Limited awareness of issues - Reactive compliance to regulations - Minimum effort when required - Rely on outside skills	- Voluntarily exceed regulations - Sporadic efforts and results - Some public disclosure - Some inside skills	- Proactive company-wide sustainability program - Triple value-based business model - Whole – system resource productivity principles - Board level oversight - Corporate Sustainability reports	- Vigorous global environmental and social development - Set leading standards and performance that others follow -Values embedded in corporate culture - Sustainability advocate

Fig. 2. Corporate Sustainability: Capability Maturity Model by Spectrum Innovation Group © 2008

[1] MSHA (Mine Safety and Health Administration) is an agency of the United States Department of Labor which administers the provisions of the Federal Mine Safety and Health Act of 1977 (Mine Act) to enforce compliance with mandatory safety and health standards as a means to eliminate fatal accidents, to reduce the frequency and severity of nonfatal accidents, to minimize health hazards, and to promote improved safety and health conditions in the nation's mines.

2.2 Co-create and Enable Sustainability Best Practices

From many companies' perspectives sustainability is largely undefined and there are few or limited standard guidelines or best practices. Our role as user researchers in this case is different from other areas of enterprise business like Finance, HR etc. that have well established best practices and business processes. As user researchers we collect data from multiple customers, industries and countries and then we are tasked with synthesizing all this information to create a consolidated view of standard sustainability practices. These new sustainability best practices may be shared with customers to be included in their sustainability strategies and programs. These best practices can also be translated into designs and concepts that are then incorporated into our products supporting customers to implement them in their business operations.

2.3 Help Modify or Define New Roles and Responsibilities

In helping define what sustainability is, we may come up with new roles and responsibilities that a company never had, or redefine existing ones. When talking with manufacturing customers about Energy Management we learned that the role(s) performing energy management related tasks in a company comes from a wide variety of different organizational departments for example Operations, Facilities, and Maintenance etc. In cases where a dedicated person is not responsible for energy management monitoring and analysis, different people from different departments are assigned the tasks. As in the case of an Energy Manager at a multi-national chemical company said, "I spend the first two weeks of the month just creating reports. My job is to focus on improvement related activities instead of worrying about creating reports!" Through research, we were able to identify areas that could be easily automated and also integrated with our existing solution. This would enable him to move from monitoring data and creating reports to focusing his efforts on strategic energy reduction programs while an *Energy Analyst* could assume the responsibility of monitoring, analyzing and reporting energy consumption.

On a recent site visit to one of our medical devices manufacturing customer, we were told that RoHS[2] regulations will begin applying to medical devices in 2014 under a proposed recast. The customer was struggling to include all the restricted substances in their current database and update other business processes to comply with the new regulation. They were also unaware of the *Compliance Specialist* role that seems fairly common in industries where such directives apply. We were able to share our experiences of talking to other customers, how the role of compliance specialists fits in their organizations, and what job responsibilities might entail. This provided our medical devices customer a deeper understanding of such a job profile and they were willing to consider including such a role in their organization in the future.

---

[2] Restriction of Hazardous Substances (RoHS) is a directive that was officially adopted in July 2006 by the European Union, for the purpose of protecting both people and the environment from hazardous chemicals found in both electronics and electrical products.

2.4 Push the Technology Envelope

Sustainability can drive innovation and we have seen examples of this in recent years with hybrid cars, Ford SmartGauge with EcoGuide, UbiGreen [5] etc. User researchers are uniquely positioned to help push the technology envelope. We can collect data that inspires developers and technologists to think “outside the box” and come up with novel solutions. As an example, during an interview with a Safety Manager at a mining customer he stated, “Aggregation of safety information across US mining industry could help in discussions around safety by driving clarity. There is a potential cross-industry benchmarking opportunity which would allow mines and MSHA to benchmark themselves against the industry to locate unsafe operations more effectively.” Such a combined view of data doesn’t exist today. This along with other data was turned into user needs and requirements for our development team. Developers now have the opportunity to create visionary technology to make this happen, leading to successful adoption of the product.

2.5 Enable Creation of Comprehensive Solutions

Sustainability is a big umbrella with many underlying topics, and this is true with our sustainability suite of products as well. Following our multi-country environmental health and safety research, we found underlying linkages with completely unrelated products that had not been considered before. The finding led to a major integration project that fostered collaboration between different teams within our organization. This laid the foundation for shared innovative thinking resulting in a holistic product for our customers.

3 Summary

The nascent field of sustainability offers user researchers numerous opportunities to practice our work in exciting new ways and makes it possible for us to influence its establishment in the future.

References

1. Brundtland, G. (ed.): Our common future: The World Commission on Environment and Development. Oxford University Press, Oxford (1987)
2. Weybrect, G.: Top 10 trends in sustainable business (2010), http://blogs.reuters.com/environment/2010/04/28/top-10-trends-in-sustainable-business/
3. Economist Intelligence Unit. Doing good: Business and the sustainability challenge (2008), http://graphics.eiu.com/upload/Sustainability_allsponsors.pdf
4. Nidumolu, R., Prahalad, C.K., Rangaswami, M.R.: Why Sustainability is Now the Key Driver of Innovation. Harvard Business Review 87(9) (2009)
5. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J.A.: UbiGreen: Investigating a mobile tool for tracking and supporting green transportation habits. In: Proceedings of the 27th international Conference on Human Factors in Computing Systems (Boston, MA, USA). CHI 2009, pp. 1043–1052. ACM, New York (2009), http://doi.acm.org/10.1145/1518701.1518861

Leading International UX Research Projects

Jakob Biesterfeldt[1] and Miranda Capra[2]

[1] User Interface Design, Truderinger Strasse 330,
81825 Munich, Germany
Jakob.Biesterfeldt@uid.com
[2] HumanCentric, 200 MacKenan Dr,
27511 Cary, NC, USA
mcapra@humancentric.com

Abstract. In a global marketplace, UX research often needs to span multiple target markets to ensure usable products for all users. While there is some literature on methods for international UX research, less attention has been paid to practical aspects of planning and conducting multi-cultural research studies. In this paper, we provide suggestions for leading international research based on our experience as part of a network of UX research companies spanning 30 countries and 5 continents. We address all stages of a typical project lifecycle, from planning and conducting research to analysis and interpretation of insights, with a focus on formative usability testing projects.

Keywords: user experience, international research, usability testing.

1 Introduction

In our experience, research which is spread among multiple companies, countries and time zones, entails higher risks than at home. Leading cooperative international research projects can seem daunting at first, but with proper planning and the right local vendor they can run just as smoothly as projects within your own country. Our goal in this paper is to provide practical suggestions based on our experience as long-term members of the International Usability Partners (IUP), an international partnership of UX consultancies from around the world. Our projects tend toward user research and usability testing for consumer and medical products, software, websites, and consumer/business electronics, generally as an outside consultant brought in either for the specific project or as part of a larger product development effort. However, many of these tips will also be helpful to in-house usability teams and international projects for a wider array of user experience research. The project stages as described in this paper are:

- Plan the Research
  - Select a project management approach (centralized, decentralized, cooperative)
  - Decide whether, when, and how you want to observe the fieldwork
  - Find local vendors
  - Collect quotes from vendors

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 368–377, 2011.
© Springer-Verlag Berlin Heidelberg 2011

- Design the research
  - Project Timeline
  - Participant Profiles
  - Study Materials
- Fieldwork
  - Recruiting
  - Planning for no-shows
  - Prepare the local team
- Analysis & Reporting

2 Planning the Research

Early and thorough planning is essential for international projects. We assume that you or your client already know the countries you would like to test in; for guidelines in selecting countries, see [3]. The next step in planning is to decide on your management approach, which will affect the type of vendor you select for the project and the scope of the bid request you send out to vendors.

2.1 Project Management Approach

Typical approaches to project management are *centralized*, where the project lead moderates in all countries; *decentralized*, where each country takes control of its own approach and study design; and *cooperative*, which is a balanced approach where the lead provides the project goals and approach, but the individual counties help craft the specific approach [5]. A centralized approach is appropriate for the early stages of design where the lead needs intimate knowledge of all findings. Examples of such situations include contextual inquiry and ethnographic research, such as in the creation of personas, market segmentation or usage scenarios [6]. A centralized approach is also suitable for studying products or services that require the project team to spend a lot of time on acquiring specific domain knowledge. A centralized approach generally incurs high expenses for the moderator to travel and hire simultaneous interpreters for all sessions, but it is sometimes less expensive for the project team to travel than to brief and debrief with all vendors on complex and unfamiliar products and services. A centralized approach allows the lead moderator to monitor sessions and stay more involved in all aspects of the research [3], but it also extends the project timeline, since you cannot run multiple countries in parallel.

We use a cooperative approach for most of our projects, with a central lead and common research approach but local researchers and local participants. Local researchers are essential because experts in the local culture, language and usage context provide the most useful insight in the local user experience. When the local moderator is fully engaged in the project goals, they are able to keep the project on track when the unexpected occurs, and can write sections of the final report, providing the voice of the local country and an additional perspective on the product being evaluated. Cultural aspects are easily overlooked or misunderstood if you are not already deeply familiar with them. This includes cultural aspects related to a country or region, but also corporate cultures or specific industry cultures. Involving multiple evaluators also results in more usability problems found than with a single evaluator [2,7], even above and beyond the benefits of involving more users [9]. We have developed a network of quality vendors in dozens of countries which facilitates a cooperative approach, and we discuss this more in the section about selecting a vendor.

2.2 Observing Fieldwork in Collaborative Studies

Observing sessions in person can be quite costly for travel. In general, the more complex the product and research goals, the more efficient it is for the research lead to travel to each vendor and observe the live fieldwork. For less complex products and research goals, the effort spent consolidating results from various countries is less costly than traveling. Simultaneous interpreters are often as expensive as UX consultants, especially if you have unusual language combinations, and interpreters may not understand technical terms used during the sessions. A full transcript and translation loses the context of emotions and attitudes. For most studies, we will watch a few sessions using remote video streaming (webcams for product testing and screen sharing for software testing), and rely on notes taken by the moderator or a note-taker to capture the key ideas from the remaining sessions. Once we have worked with a vendor a few times, trust their expertise, and work using methods and materials familiar to both the project owner and the vendors, we don't need to observe any of the sessions. Decide how you want to observe either in person, live-remote or not at all before sending the project out to bid, because it can affect the type of room needed, days to run the study, time of day for sessions, etc.

2.3 Finding Vendors

In a cooperative model, it is most critical to partner with the right local vendor, someone that you trust because you will be relying heavily on that local vendor for both the quality of the research and understanding of the local culture. Allow plenty of time to find, contact, exchange costs and other information as well as securing availability with vendors, if it is a first contact. They need to have excellent communication skills to ensure smooth project planning, clear discussion of issues that arise, and a proper write-up of their section of the final report. Vendors must have a common understanding of the methods, processes and scientific background in UX so that they can answer your research questions and challenge your methodology, if needed. They should be experienced in international projects and understanding cultural differences, and have domain expertise in related products or software.

To find a qualified vendor, you can consult networks of companies that specialize in international projects, such as our network, the International Usability Partners (IUP), online directories from groups like UPA and SIGCHI, universities and market research companies. Be careful using companies that specialize in market research; while they will have smooth project management and recruiting, they may not have a background in usability for moderating tasks, localizing tasks and questionnaires, and other UX-specific activities. Similarly, be careful of companies that say they have usability experience but it are purely survey-based, and do not have a background in behavioral research [3]. Colleagues and mailing lists or discussion groups of professionals and academics in your field are also excellent sources.

We recommend establishing long-term relationships or using networks that have already established long-term partnerships. It not only reduces the effort involved in finding a vendor, it will also give you peace of mind with regards to the quality of their work and you can continually improve the collaboration until it feels like you're working with your own company

2.4 Bidding Process

When soliciting bids, create a detailed project specification and set of questions that you will provide to every vendor, and send a quote template to make sure you get comparable numbers from each at the level of detail you require. Specificity in this initial bid request is especially important when working across multiple time zones because time delays make it difficult to have back-and-forth discussions. Be prepared to stay up late or get up early to catch the vendor when they are at work. It helps to involve all vendors in each Q&A, i.e. send questions from one vendor and your answer to it to all other vendors, too, so that everyone has the same understanding of the project scope and approach. Table 1 provides a list of items that we review when creating our bid requests; we select the ones that are appropriate for each study. Make sure you specify the currency for the bid, and terms for handling changes in exchange rates between the time of bidding and project invoicing. Be prepared to receive quotes ranging from 0.5 to 3 times as much as you would calculate at home. European and North American vendors have somewhat similar costs; China and India are at the lower end of the range, with Japan on the upper end.

Table 1. Checklist of bid request items

Common Items	Occasional Items	Other information
Preparation	Simultaneous translator	Moderator bios
Moderation	Transcript	International experience
Note-taker	Participant food	Company profile
Video recording	Observer food	Facility photos
Topline report	Check fee	Recommended hotels
Full report	Parking fee	Local holidays
Recruiting	Video streaming	Days of fieldwork
Incentives	Printing	Fieldwork schedule
Facility fees	Shipping	Payment schedule
Video recording	Import taxes	Bid currency
Translate study materials	Suggested over-recruit	No-show policy
Moderator debrief		
Moderator notes		

Before confirming vendor selection, schedule a live phone call with the vendor, preferably with the specific moderator for the study. This in-person communication will help validate that you can have effective conversations without difficulties due to language barriers. This is also a chance to determine if you have similar approaches to study design and report writing, and if the moderator has sufficient familiarity with relevant techniques and standards.

3 Designing the Research

This section will discuss how to create a project plan that will translate well to other countries and for other facilitators.

3.1 Project Timeline

We prefer to run in the lead country first, to smooth out any issues with the study design and make the materials as complete and detailed as possible before providing them to the local moderators. The local moderators can then run in parallel, unless there are restrictions such as the client wanting to watch all sessions live or limited availability of product prototypes. If the timeline is too short to run all sessions in the lead country before starting the other countries, run at least 3-4 sessions to allow sufficient time to smooth out the worst issues with the protocol. The session schedule will depend on many factors, including:

- Public holidays, school schedules and popular vacation weeks in each country
- Cultural differences in the acceptability of daytime, evening and weekend sessions
- No-show rates for the local culture and specific facility
- Flexibility of dates when the vendor has their own facility versus tight schedules when using a rented facility

In China, for example, most testing is done on the weekend and it can be impossible to find any participants during working hours. In Germany, the weekends are usually kept free for family life, but it is easy to find participants during the week. In France, most recruiting agencies are closed in August.

3.2 Screeners and Participant Profiles

Creating a screener to get the proper people is always a painstaking process, with careful selection of criteria and phrasing of screener questions. It is tempting to maintain strict control over the screener to keep the user populations as similar as possible from one country to another, but this is impossible because user populations differ from one country to another and you may need different ways of identifying target users [3]. There may be questions such as cultural class, income, and life style which often cannot be directly translated from one culture to another [4]. Describe the target group in familiar terms to local moderators and participants. Make sure the target population (or its equivalent) exists and the label given to it makes sense in the target country. Explicitly describing the type of person required is often more useful than specifying exact figures (e.g. ‘holds a senior managerial position’ instead of ‘earns at least $80,000 a year’) [4]. Rely on the local facilitator to suggest changes as well as doing some preliminary research on user groups and contexts of use, especially when your research requires specific domain knowledge. For example, a while ago, we did some research on software for hearing aids. We found several differences between Germany, the UK and the US that were good to know during the recruiting process. We had to find participants who would cover the usage of a wide variety of health insurance options and procedures, completely different in all three countries. Typical training for current users of the software ranged from a 3 week course with one of the manufacturers (US) to a 3 year training for German users to highly qualified doctors of audiology in all countries. Often the screener feedback from the local facilitator is helpful in understanding differences in the product users from one country to another, and may be interesting to provide to your client or product team as a project learning separate from the other findings of the user research.

3.3 Study Materials

Explain the research goals in the study guidelines: What are the reasons for choosing this method / test setting / scenario? This helps vendors to understand the project scope. Make sure the vendor gives feedback on any localization issues regarding the scenarios. Adjust the scenarios / session guide if necessary. Provide clear guidelines on how to answer the research questions. Make sure the guidelines as well as products/prototypes under test are translated correctly [1]. Although fieldwork is done in the local language, all deliverables and communication is usually done in English. One way of ensuring correct translation of critical information is to have the content in question translated back to the original language by an external translator.

4 Fieldwork

This section will discuss issues related to ensuring that the fieldwork runs smoothly.

4.1 Recruiting

Practical recruiting issues include getting participant consent, incentivizing participants, dealing with “no-shows” [4].

Video and Data Collection Consent – laws safeguarding personal information vary from one country to another; your local vendor should be able to help you

Difficulties recruiting – if you have having trouble finding a particular user group, there are many possible reasons: the local vendor misunderstood the screener, the screener has a logical flaw in it, your local vendor has an inadequate database, or the user population that you wanted does not exist in the local country. The last point is something that needs to be carefully considered in international research. Is it possible that the product team does not properly understand the local market? This is a good time to have a call with the local moderator to understand why they think they are having a hard time recruiting. Best practice among our partners is to invite recruiters to a usability lab, show them around and make sure a mutual understanding is established of the needs and pitfalls unique to usability testing, as compared to market or survey research. It helps if the recruiter has understood the relevance of having participants who have an opinion that they are willing and able to get across [4].

Incentives - The incentives required to persuade potential participants to take part vary a great deal from country to country. This is linked to the cost of living - incentives required in the UK are substantially higher than in most of mainland Europe [4]. In China, more so than the US or Europe, there is a delicate balance to be struck in offering the correct incentive. The smaller the target segment, the higher the incentive required. Incentives must be very high to persuade people of high social status to participate. Testing during working hours also increases the figure required to persuade people to take part. However, care must be taken when offering above market-rate incentives, as we have found that many Chinese people can be suspicious of being offered obviously inflated sums [4].

4.2 Planning for No-Shows

No matter how many times we call or email participants, there will always be some that do not show up for their appointment. We assume a 10% to 30% no-show rate for a study. The specific rate depends on how much you are paying participants, the target market, cultural differences, and the characteristics of the specific recruiting agency. No-shows are an issue for any user research, but even more so for international projects because customs and conventions may differ from one country to another, and multi-country projects can be complex while still having tight timelines. In international projects, often more people are involved in each session than in local projects. You don't want observing clients, note takers, project owners or interpreters to waste time waiting for no-shows. However, if it does happen, your clients will be grateful if you have prepared something like showing a video recording of a previous session or present some related findings from your experience, or an agenda for a working meeting related to the project (mid-project readout, planning for the next phase, etc.). If you understand the different approaches to handling no-shows, you can work with your local vendor to select one that is best for the specific project.

- **Recruiting extra participants for all sessions** – this works best for focus groups, gang surveys and other group work, where you can recruit one extra person to cover any of a number of people in a group.
- **Recruiting floaters** – these are people that are paid to sit around for more than one session, and you use them if there is a no-show. You usually pay one recruiting fee and less than the equivalent incentive for the individual sessions covered. You risk having an empty slot if there are adjacent no-shows.
- **Adding extra sessions to the end of the study** – if there is a no-show, recruit another person at the end of the study. This approach works best if the local moderator owns their own facility (no additional fees), there is little or no travel, you have other activities for the open slots, and you are not on a tight timeframe.

Which one is more cost effective depends on the specifics of the study. Additional considerations to account for no-shows are:

- How will this affect the participant budget? Does the vendor charge you for everyone recruited (usually) or just for those that show (unusual)?
- How will this affect the facility budget? Does the vendor charge you for sessions you don't run or for a full extra day for added sessions?

- How will this affect the project timeline? Will adding extra sessions delay the project? Can you make use of empty slots for productive work or is that time lost?
- How precise is your recruit? Do you have some flexibility in replacing a no-show from one group with someone from another group?
- How tough is your recruit? Will it be hard to find replacements at the last minute?

4.3 Preparing the Local Moderator

In user research projects, using multiple moderators can create issues because there is an evaluator effect, with different moderators uncovering different usability problems [7], even when the study methodology and participants are identical [2]. When working with a new vendor, it is also difficult to get a full understanding of the quality of the evaluation and report during the planning phases of the project [1]. Through experience on many projects, we have developed a set of best practices to minimize issues and take advantage of the expertise of the multiple moderators on a project.

- Run the study in the lead country first (or several sessions, if time is extremely tight) so that you can provide the local moderator with a detailed guide that addresses common mishaps and misunderstandings that happen during the session.
- We cannot emphasize enough that the local moderator needs to understand the study goals so that they can interpret the moderator guide, handle unexpected situations, filter the most important findings from the sessions, and know how to adjust the protocol if a participant is having difficulty or is running out of time. Make sure they understand the reasons for choosing this test method and scenarios, the project scope.
- Have a live conference call to brief the local moderator. This is essential to ensure that the project goals are properly communicated and all questions are answered.
- Provide a video recording of a few sessions from the lead country for the local moderator to watch. Avoid the first few sessions, where the lead moderator is still smoothing out the study protocol.
- Provide the local moderator with a first categorization of insights to look out for. UX research tends to be explorative, uncovering issues as you go along, but providing high-level guidance will help maintain consistency and insure quality.
- Provide a detailed presentation template in order to integrate and compare results efficiently. Provide this template before the local moderator begins fieldwork, to make sure that all report topics are addressed in the user sessions.
- Ask the local moderators for feedback on localization issues and methodological concerns.

5 Analysis

Challenges in collating results across countries include

- Getting all results from the vendor to the client
- Identifying cultural differences and commonalities
- Completing findings in all countries, cross-checking against other markets.

We use the following strategies for getting the most out of our projects.

- The local moderator writes their section of the report. There are so many nuances in a usability study that may not come across fully in the note-takers notes or videos, so you want to take full advantage of the moderator's expertise. We also rely on the local moderator to provide their expert opinion of the interface, above and beyond merely reporting what they explicitly observe users to do. Expert reviews benefit from multiple evaluators and result in more problems found [8].
- Provide a report template for the local moderator before the fieldwork starts. This will help ensure that the moderator makes the proper observations.
- A minimalist approach is to have the local moderator write an executive summary of high-level take-aways, and then supplement with details from the note-taker's notes.
- A debriefing call, preferably with all markets at once

The debriefing call is a critical part of gathering the conclusions from the study. These miniature "expert focus groups" provide the most valuable insights into cultural differences and commonalities. It captures many subtle issues that may be difficult for a foreign moderator to explain in a written report but will come across better in conversations between moderators. Often we find that one of the vendors reports something that none of the other vendors did. During the call we can try to understand whether this is due to differences in the type of participants, cultural issues, moderation style and the evaluator effect [2,7], usage habits, etc. Sometimes we find that other moderators observed the issue but did not note it in their report. Having a conference call with all facilitators is essential to resolving these issues.

6 Additional Topics

There are many small details that can create large difficulties if left unattended.

- **Customs** – make sure you allow several extra days in shipping for prototypes to clear customs, be clear who is paying duties/taxes, and avoid shipping food
- **Technical standards** – provide enough plug adapters for all equipment; provide a power converter if necessary; make sure you can view the local video recording format (NTSC vs. PAL, regional DVD encodings, etc.).
  - Mobile communication standards and SIM cards can vary from one country to another, if your study involves mobile phones
- **NDA** – have a standard one available in several languages, including English
- **City for Fieldwork** – you can frequently save costs by allowing the local vendor to pick the most affordable city; usability is less subject to differences in regional markets than traditional market research questions about tastes and pricing expectations. However, some clients will balk at accepting research from unfamiliar markets, and it is sometimes worth running in a city with name recognition. If your clients are planning to travel to observe, it may be cheaper for them to travel to a large city, and easier to be in a city where they do not have to rent a car.

- **Breakages and debugging** – when working with prototypes, make sure there is local talent available to debug software compatibility, solder broken wires, etc.
- **Remote observation** – there are many software tools that are readily available now for streaming webcams and sharing desktops, which makes the decision not to travel easier, while still allowing monitoring of remote sessions, when necessary
- **Cultural challenges** – will the participants accept female facilitators? Do they say "no" when they mean "no"? For which tasks do the users use the product in various countries? Is the session guide localized? Are users reluctant to be critical about a product?
- **Internet access** – do users have access to the site you want to test? Do you understand common websites and applications that they are familiar with? Is their internet access censored?

Acknowledgments. We would like to thank our partners in and around the International Usability Partners for their continuous support and input for this paper.

References

1. Biesterfeldt, J.: Managing International Research Vendors. In: 2008 Conference Proceedings. Bloomingdale, IL, UPA (2008)
2. Capra, M.G.: Usability Problem Description and the Evaluator Effect in Usability Testing. Unpublished dissertation. Virginia Tech, Blacksburg, VA (2006), `http://scholar.lib.vt.edu/theses/available/etd-03222006-201913/`
3. Dray, S., Siegel, D.: Sunday in Shanghai, Monday in Madrid?! Key Issues and Decisions in Planning International User Studies. In: Aykin, N. (ed.) Usability and Internationalization of Information Technology. Lawrence Erlbaum, New Jersey (2005)
4. Gillham, R., Biesterfeldt, J., Thalmeier, S., Fuchs, F., Shang, T., Sabena, E.: Guidelines for Successful Recruitment in International Usability Studies. In: Valone, D., Page, B. (eds.) 2007 Conference Proceedings CD, Bloomingdale, UPA (2007)
5. Gorlenko, L., Krause, S.: Managing International Usability Projects: Cooperative Strategy. In: Grinter, R., Rodden, T., Aoki, P.l., Cutrell, E., Jeffries, R. (eds.) Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, pp. 159–164. ACM, New York (2006)
6. Mack, A., Dray, S., Larvie, P., Lovejoy, T., Prabhu, G., Sturm, C.: Managing International User Research. In: Grinter, R., Rodden, T., Aoki, P.l., Cutrell, E., Jeffries, R., Olson, G. (eds.) Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, pp. 5–8. ACM, New York (2006)
7. Molich, R., Dumas, J.S.: Comparative Usability Evaluation (CUE-4). Behaviour & Information Technology 27(3), 263–281 (2008)
8. Nielsen, J.: Heuristic Evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–62. Wiley and Sons, New York (1994)
9. Woolrych, A., Cockton, G.: Why and When Five Test Users aren't Enough. In: Vanderdonckt, J., Blandford, A., Derycke, A. (eds.) Proceedings of the People and Computers XV - Interaction without Frontiers: Joint Proceedings of HCI 2001 and IHM 2001, pp. 105–108. Cépadèus Éditions, Toulouse, France (2001)

Lessons Learned from Using Personas and Scenarios for Requirements Specification of Next-Generation Industrial Robots

Petra Björndal[1], Mikko J. Rissanen[1], and Steve Murphy[2]

[1] Industrial Software Systems, ABB Corporate Research
Forskargränd 7, 721 78 Västerås, Sweden
[2] ABB Discrete Automation and Motion, Robotics,
Lunnagårdsgatan 4, 431 90 Mölndal, Sweden
{Petra.Bjorndal,Mikko.Rissanen,Steve.Murphy}@se.abb.com

Abstract. As development and productization of complex technology for new markets is challenging, it is crucial to get valid information about the intended future users and represent that information in a proper way to technical experts and project management. This case study describes experiences from using the "engaging personas and narrative scenarios" approach [1], [2] for defining requirements for next-generation industrial robots. We found that additional steps were necessary to supplement the methodology to fit the development of complex industrial robots, namely globalization, validation of personas among end users, prioritization of personas, creation of common vocabulary, identification of business critical scenarios and identification of safety critical situations. The main benefit from using personas and scenarios was their role as communication catalyst and how the descriptions facilitate building a common vision within the project team.

Keywords: Persona, Scenario, UCSD, Robot development.

1 Introduction

This paper discusses the use of the "engaging persona and narrative scenario" [1] methodology during the requirements specification phase of a development project for next-generation industrial robots at ABB. User-centered systems design (UCSD) [3] and Human-computer interaction (HCI) have been found to add value to the Human Robot Interaction (HRI) domain [4], but UCSD has not evolved into widely accepted practice for developing industrial robots. The discussion around the definition of User Experience (UX) has been ongoing for a long time in the HCI community, but Law et al. [5] found with the help of systematically gathered views on the nature and scope of UX in the HCI community that the (by then) draft definition in ISO 9241-210 [6] *"A person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service"* is agreed upon. UX is perceived as grounded in UCSD practices and part of the HCI domain and deals with different possible benefits users may derive from a product. User Experience (UX)

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 378–387, 2011.
© Springer-Verlag Berlin Heidelberg 2011

takes a broader perspective than HCI and HRI, which mainly focus on user-technology interactions. UX considers design, users, and business aspects. This perspective is used in this work.

2 The Project

The project referred to here is an implementation of a future vision of next-generation industrial robots. The goal is to develop a robot with increased flexibility and user-friendliness and reduced need for safety equipment like fences surrounding the robot. Purely technological discussion about future possibilities had been going on for several years. Now there were five large technical sub-projects, of which each had their own vision and specific technical focus (Fig.1). Early in the project a high level concept was formed, but it gave little guidance for technical development toward a robotic product. In absence of a detailed enough common vision, sub-projects formed their own visions about what features the product should have and perspectives to the most difficult technical problems. The lack of a common vision obstructed further requirements specification for the robot. After a period of technological development, a need of overall understanding of both intended users and the context of use was identified by project management. One of the sub-project leaders recognized the need for usability work to support the requirement specification process and asked for support from the usability team in ABB Corporate Research: "*You need to help the technical team to overcome the points where we are stuck.*" There was a need for a consensus about key features of the new robotics product and the consensus had to be based on solid data about the intended users.

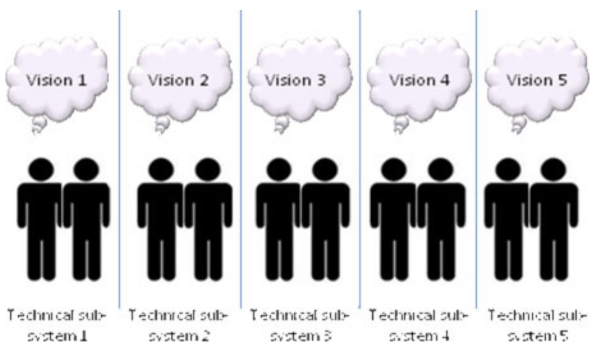


Fig. 1. Setup: Sub-projects had their own visions. Common vision was undefined and requirement specification therefore difficult

A structured way of working and the ability to facilitate discussions were desired by the project management. Based on previous own experience, and documented good effects on communication between teams and an ability to facilitate user-focused design discussions [7], the "persona" and "scenarios" methods were chosen

[1], [8], [9]. Because of intellectual property concerns, the project was kept confidential and field-trips to end-customer companies were kept to minimum. Five trips to customer sites were done, including workshops with managers to find interviewees, observations, and 21 interviews with potential end-users relevant for the project. In total, 49 person hours were spent on doing face-to-face interviews. Field-trips were done to three different countries in Europe, but trips outside of Europe were not possible within the expected schedule, though the target market for the product is the whole world. The two person usability team cooperated with 30 project members distributed over 4 countries during the project time. Another 56 project members had access to usability material through a network server.

3 Approach

The approach of the usability team in this project was based mainly on Lene Nielsen's 10 steps to Personas [2] which were extended during the work with 6 additional steps in order to provide and communicate a holistic view of the needs of the end-users (see Fig.2). Below is a stepwise and chronological description of the experiences with persona and scenario development in this technically challenging project that involved several stakeholders, technical experts, project managers, a product manager and other key persons. Steps deviating from Nielsen's description are marked with "Additional Step" which indicates the need for practical adaptations in this UCSD methodology for developing complex robotic products.

3.1 Finding the Users

The usability team used already established contacts with customers through product management, and after customers were identified and Non-Disclosure Agreements finished, workshops were arranged with customer representatives to identify relevant interviewees. Interviews and observations were done by the usability team. The field data was collected using task analysis [10]. The task analysis was complemented by interview questions to collect information about interviewees' background, personality, interests and concerns - all important data in order to be able to build engaging personas.

3.2 Building the Hypothesis

All notes were written down in a consistent way and collected on the project server together with other background data. During this process the usability team started to build hypotheses about personas to be covered and asked questions such as "are there two main types of customers, high-tech and low-tech oriented?" Much of this discussion was done collaboratively with the project management.

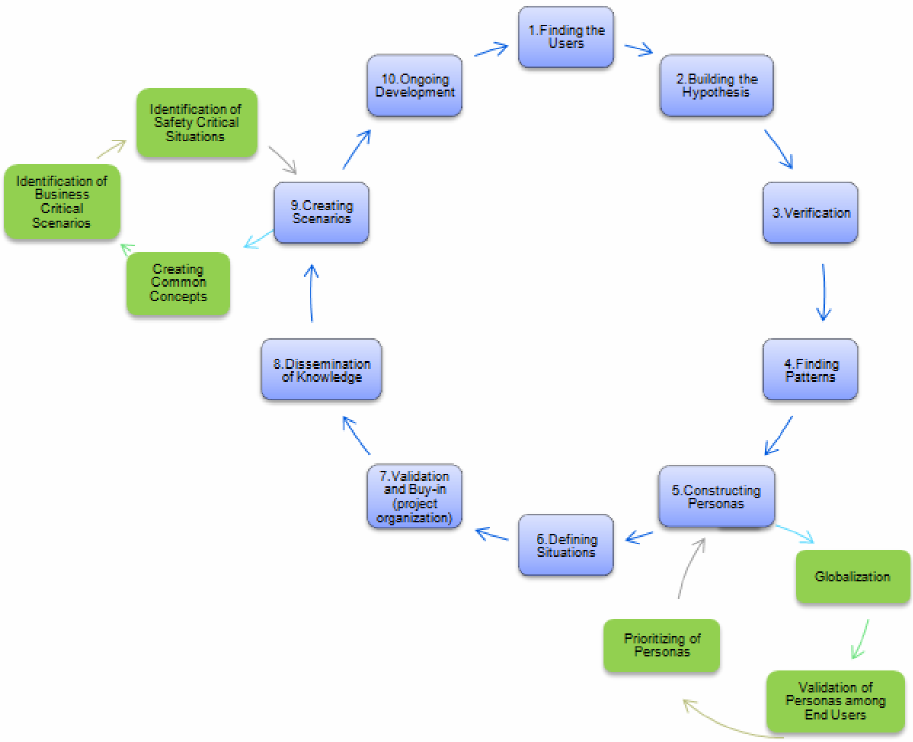


Fig. 2. Lene Nielsen's process 1-10 (blue) supplemented by additional steps necessary in this project (green)

3.3 Verification and Finding Patterns

The usability team wrote the persona descriptions in iterations. After each iteration other participants in the project reviewed the tentative persona descriptions in workshop sessions. Discussions in the workshops resulted in more detailed data about level of technical knowledge of the personas, which was supported by already collected user data. Already early in the process, the initial draft persona descriptions were well received and helped a lot to facilitate discussion, even if they had to be significantly modified later on. When the usability team presented the first version of personas, they were perceived as trustworthy with clear traceability to collected user data. All collected user data like notes from interviews and observations were easily accessible for all project members for further reference during the project.

3.4 Constructing Personas

The organization had previous experience of working with personas; the format was already known to the technical teams and therefore well accepted. The usability team took these good practices and refined them for the specific project. Photos were chosen from an internal photo database providing realistic photos resembling actual people interviewed during field trips, not models. Background, knowledge and skills of personas were described together with attitudes towards, for example, computer usage. Goals, concerns and work practices were also part of the description. The personas were restricted to maximum one A4-page length and the team took care of only adding a few sentences of fictional data about each persona's personality. Eight different persona descriptions were written in this phase. The team took great care to ensure that each persona had a very distinct personality. For example, a one-liner such as *"I speak [the programming language], Finnish and English – in that order."* The one-liners were well understood by project members: "*You get an instant feeling for the persona."* Some of the fictional background was easily memorable and got, for example, the product manager's attention: "*… and Joe's [priority 1 persona] wife complains a lot."* One-liners are part of many descriptions of personas [8], [9] but not part of Nielsen's approach and not part of persona descriptions used earlier in the same organization. The usability team found such one-liners very effective in the daily communication in the project as they gave a quick summary of what each persona's role in the customer organization is as well as their main focus. The one-liner was often used in reference to the persona description instead of its name.

3.5 Globalization (Additional Step)

When iterating the personas, comments from the project team, such as *"too much old fashioned Northern Europe feeling now,"* resulted in changes to names, photos and certain details to align the descriptions better with the vision of a product used in a global market. All other aspects were kept intact in the personas. One additional persona was written in this phase in order to cover more technical aspects of the customers' personnel, which was a response to inquiries made by several members of the technical project team. Impartial user data was supplemented with second-hand data from customer representatives less involved in the project and knowledge coming from project members. Assumptions about possibly missing data in terms of market coverage were discussed. The globalization step was necessary for credibility and important to ensure wider buy-in in the project.

3.6 Validation of Personas among End Users (Additional Step)

Validation of personas among end users which also involved the globalization aspects refined the personas iteratively and added to their trustworthiness. Workshops with some of the customers were organized, and also one workshop with relevant customer representatives not visited during fieldtrips. Comments such as "*in general, the responsibility areas [of the personas] look real*" strengthened the validity. "*Nothing big is missing*" and "*this is a guy we could hire here*" were taken as confirmations about the validity of the personas. Several details were added and some customers complained about age span of the personas. The interviewees chosen by the customers' management were usually the most experienced ones. Therefore, some of the personas did not represent a completely normal organization, which lead to a quick revision of the ages, backgrounds and skills of some of the personas. Validation among end users is discussed as a possible improvement of the process by Nielsen [1]. In this project the usability team found it necessary to prove validity of the descriptions in order to get full acceptance for them in the project.

3.7 Prioritizing of Personas (Additional Step)

To facilitate the building of a common vision in the project, prioritization of personas was proposed by the usability team. Prioritization was done using an effect map [11] onto which the personas were placed (see Fig.3). The effect map showed the overall expected effect of doing the project, the prioritized personas, their respective main needs, and was later extended by links to solution descriptions described in use cases. Thus, for the first time in the project, the users' needs were presented to the project management and the product manager and discussed with them in a structured way. This resulted in an official prioritization for future situations requiring certain trade-offs. This was an important step towards consensus on the main focus for the product in the product management. "*Joe [priority 1 persona] is the sledgehammer,*" said a sub-project leader to stress the importance of persona *"Joe"* when discussing argumentation for certain technical requirements that might increase the budget.

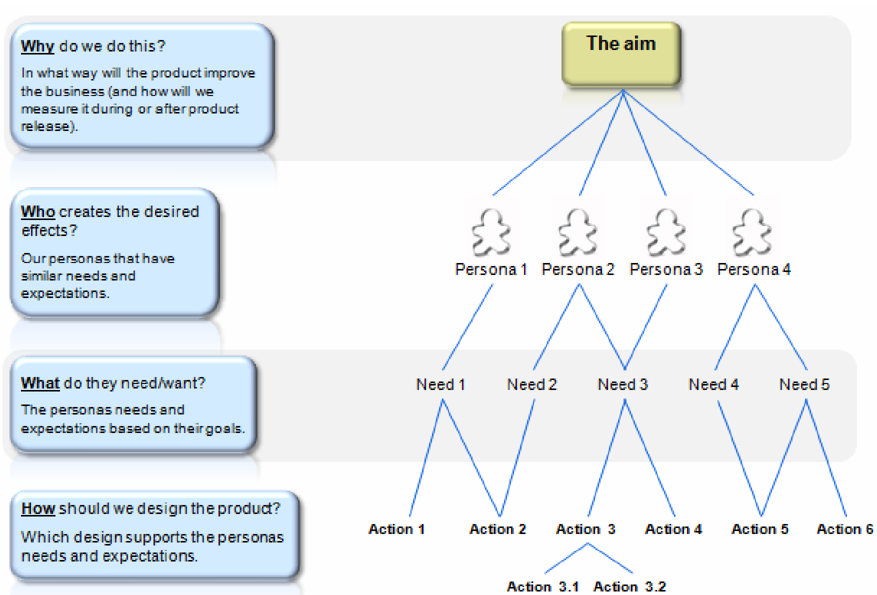


Fig. 3. A schematic picture of an effect map as explained to the project management

3.8 Defining Situations

The persona set, their tasks, goals, motivations, and general context were the basis to find interesting topics for scenarios. The definition of scope of work to support in the future was discussed in the project. The usability team was responsible for defining the scenarios, but also dependent on inputs from project members on certain technical issues.

3.9 Validation and Buy-in

The usability team hosted a series of workshops to discuss, present and prioritize persona descriptions and scenarios with the aim to involve stakeholders and to communicate results and insights. A good overview presentation of persona and scenario material played an important role in getting involvement from the technical sub-projects and to be able to facilitate the ongoing discussions among stakeholders. A brief process description of the work enhanced understanding and gave an overview of the work.

3.10 Dissemination of Knowledge

As a part of the workshops, project members were informed on how to use personas and scenarios and also on how to link this material to use cases - the next step in the development process. All data from interview notes to presentations were accessible for the project member through the project server.

3.11 Creating Common Vocabulary (Additional Step)

Before starting the scenario writing process, the usability team compiled fundamental concepts and naming conventions together with project members. It was necessary to form a common language to be able to write meaningful, credible scenarios without serious misinterpretations. The common terminology formed a “project language” and thus contributed considerably to building a common vision in the project. The concepts were based on basic assumptions about the robot such as the viewer’s reflection angle that determines, for example, which side is meant by “left” and “right” of the robot. This information was not written down in any other documentation common for the whole project and therefore necessary before writing scenarios.

3.12 Creating Scenarios

Topics for scenarios were found in collaboration with other project members and based on earlier discussed outlines of personas using the new robot. Known problematic situations were also documented. The necessity to identify and describe only the most business-critical scenarios was recognized, due to the complexity of the project.

3.13 Identification of Business Critical Scenarios (Additional Step)

The 23 most critical scenarios were selected together with stakeholders to be the basis for the major requirements. The usability team was responsible for writing the scenarios with support from other parts of the project. The scenario approach challenged the different visions in the project and scenario descriptions unveiled missing functionality. The holistic approach of the scenarios and the possibility to add earlier collected field data created detailed stories with a lot of context information. All scenarios were descriptions of one or several personas in a specific context with specific goals. Not only straightforward cases were described, but also problem situations like a technically demanding troubleshooting scenario, authored by one of the sub-project leaders.

3.14 Identification of Safety Critical Situations (Additional Step)

The goal of the project, to develop a robot with reduced need for safety equipment, put special importance on the topic of safety. The team quickly realized that the overall view provided by the UCSD process facilitates the identification of safety critical situations. Working on safety aspects was originally not planned for the usability team, but became necessary later when the team had more information about customers' work practices and everyday working situations. In this way, the user-centered way of thinking supplemented existing technical expertise and helped to identify and fill gaps between the risk assessments done in different sub-projects For instance, the frequency of occurrence of each situation in the formal risk assessment was based on input from the task analysis and persona descriptions.

3.15 Ongoing Development

The responsibility for persona descriptions and scenarios, including any changes to them, was kept with the usability team when the project entered the phase of writing use case descriptions. Personas and scenarios played a key role to provide context to use cases and also affected other discussions during later phases of the project.

4 Impact on Design Discussions

The main impact of personas and scenarios was their role as communication catalyst which enabled building a common vision from the different sub-system-oriented visions (Fig.4). The process to put personas and scenarios on paper created important discussions and the descriptions made complex relations easier to talk about – in a language common for all stakeholders, from end user representatives to the product management. As a result, knowledge about the needs and goals of the targeted users were well known and influenced the project work on a day to day basis. Everyone talked about *"Joe"* and his needs for the most challenging technical solutions. As a consequence, the requirements for the robot were for the first time discussed in a structured way in relation to valid user data, which helped technical experts clarifying the vision and then breaking it down into practice. As discussions got more structured, the complexity of the product became evident, missing functionality was identified and requirements were reprioritized.

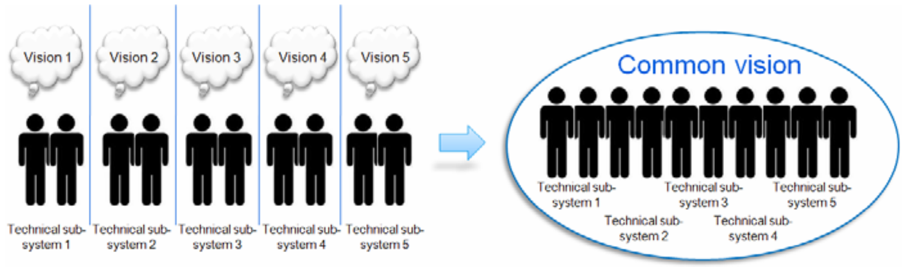


Fig. 4. Result: Common vision enabled by discussion about personas and scenarios. Requirements specification became easier

5 Conclusion

Because development and productization of complex technology for new markets is challenging, it is very crucial to get valid information about the intended future users early and to represent that information in a proper way to the technical project team and the management responsible for the new product. In this project we followed the "engaging personas and narrative scenarios" process [2] for developing next-generation industrial robots and found it necessary to extend existing 10 steps with 6 new ones, driven by real-world requirements and constraints.

- *Globalization* to align persona descriptions to a global market when only part of the markets were properly covered by field trips and user studies,
- *Validation among End Users* to enhance the credibility of the descriptions and facilitate buy-in from technical experts,
- *Prioritizing of Personas* to ensure user-centeredness in the minds of project management,
- *Creating Common Vocabulary* to avoid misinterpretations,
- *Identification of Business Critical Scenarios* to keep focus on the most important scenarios to the customers and to the developers, and
- *Identification of Safety Critical Situations* to facilitate formal safety risk assessment.

Future research will study if these steps have wider applicability in UCSD of technically complex products.

Acknowledgments. We thank Magnus Larsson, Martin Naedele, Mats Källman, Henrik Jerregård and Björn Matthias from ABB Corporate Research for their support.

References

1. Nielsen, L.: Engaging Personas and Narrative Scenarios – a study on how a user-centered approach influenced the perception of the design process in the e-business group at AstraZeneca. PhD Dissertation. Department of Informatics, Copenhagen Business School, Denmark (2004)

2. Nielsen, L. (Guest Author): Ten Steps to Personas, `http://www.hceye.org/HCInsight-Nielsen.htm`
3. Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., Cajander, Å.: Key Principles for User-Centered Systems Design. Behavior and Information Technology 22(6) (2003)
4. Adams, J.A.: Critical Considerations for Human-Robot Interface Development. AAAI Technical Report FS-02-03 (2002)
5. Law, E.L.-C., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J.: Understanding, Scoping and Defining User eXperience: A Survey Approach. In: CHI 2009 (2009)
6. ISO 9241-210:2010(E) Ergonomics of human–system interaction — Part 210: Human-centered design for interactive systems, page 9. International Organization for Standardization (ISO), Switzerland (2010)
7. Long, F.: Real or Imaginary; The effectiveness of using personas in product design. In: Proceedings of the Irish Ergonomics Society Annual Conference, Dublin, pp.1-10 (2009)
8. Pruitt, J., Adlin, T.: The Persona Lifecycle: Keeping People in Mind Throughout Product Design. Morgan Kaufmann, San Francisco (2006)
9. Cooper, A.: The Inmates are Running the Asylum. Macmillan, New York (1999)
10. Hoffman, R.R., Crandall, B., Klein, G.: Protocols for Cognitive Task Analysis. In: Working Minds: A Practitioner's Guide to Cognitive Task Analysis, MIT Press, Cambridge (2005)
11. Ottersten, I., Balic, M.: Effect managing IT. Copenhagen Business School Press, Copenhagen (2007)

A Multi-tiered Approach to Enterprise Support Services*

Coimbatore S. Chandersekaran and William R. Simpson

Institute for Defense Analyses, 4850 Mark Center Dr., Alexandria, Virginia 22311

Abstract. The Enterprise Support Desk (ESD) is the combination of people, hardware, deployed software agents, and software displays, which maintain the health of the enterprise service based operations. It is both pro-active and re-active. It is required to be integrated with hardware and software health monitoring systems deployed by the enterprise services provider. The objective of this paper is to provide the basic architecture being employed by the USAF enterprise system.

Keywords: Help Desk, enterprise, support services, information sharing.

1 Introduction

The Enterprise Support Desk (ESD) is the combination of people, hardware, deployed software agents, and software displays, which maintain the health of the enterprise SOA operations. It is both pro-active and re-active. It is required to be integrated with hardware and software health monitoring systems deployed by the Area Processing Center (APC) or enterprise services provider. The ESD attempts to resolve 90% of customer issues within 20 minutes, and average less than 15 minutes resolution for all issues. The ESD in this case consists of four basic levels and three separate groups:

1. Level 0: Client Self-help. Assistance is provided to the client in the form of Knowledge Repository (KR) access, frequently asked questions (FAQ), and diagnostic software. Failure to resolve issues at this level leads to a call to the help desk.
2. Level 1 - customer support and help desk. A client may ask for assistance through a help desk phone number, e-mail, or web form entry. This is the unit that will develop a Trouble Report (TR) and see most through to completion.
3. Level 2 – Proactive monitoring of services. This unit will monitor network activities and the performance of services using a testing tool, and a series of embedded agents. This level will assist level 1 when resolution of help desk requests has not been completed at that level. It may also generate TRs before clients call to report problems.

---

* The publication of this paper does not indicate endorsement by the Department of Defense or IDA, nor should the contents be construed as reflecting the official position of these organizations.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 388–397, 2011.
© Springer-Verlag Berlin Heidelberg 2011

4. Level 3 – Active monitoring and security. This unit will monitor all security alerts sent by agents and based upon heuristics developed within the enterprise will perform remote desk audits and dispatch a security monitoring team for physical audit of indicated desktops. This level will assist level 1 and level 2 when resolution of help desk requests have not been able to resolve the issue.

In order to facilitate all three units a series of agents on desktops and service machines, a knowledge base system, display work stations, and alert policies are needed. These requirements are varied and must be integrated with hardware and operating system health monitoring data. For this reason, all three units will use service monitoring software, which is flexible, configurable, expandable, and adaptable to include information from health state monitoring activities. Further, every desktop or laptop within the Air Force Enterprise will have embedded agents for self help and repair of common problems as well as software (TBD) that will allow ESD principals to take control of the hardware unit for the purpose of auditing hardware and software configurations and provide troubleshooting assistance. Figure 1 shows the architecture of agents for the services management. There are separate agents for hardware health monitoring and two types of agents for each service. The first agent is embedded in the service itself for provision of alerts and internal logging of service data. The second agent is installed on the server and provides a sweep of log files, either periodically or on demand. At least one vendor (Amberpoint) [21] provides both such agents.

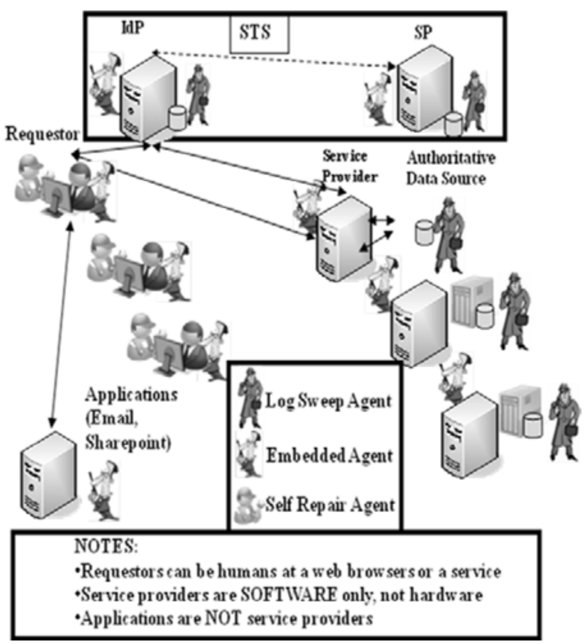


Fig. 1. Deployed Agent Architecture

Each Unit will have an administrator present with specific duties. Each shift will have such an administrator and the team of administrators will meet at a frequency dictated by events to review operations and to modify or create heuristics for ESD usage. Administrator privileges are required for certain tasks. All ESD personnel must abide by enterprise security policies, including bi-lateral authentication and SAML authorization for access control. Figure 2 illustrates how these units work with the monitored data.

2 The Knowledge Repository

The knowledge repository (KR) is a single integrated source of all information on the operation of the enterprise. It will be updated by all three units within the ESD and accessible to all three units within the ESD. Instrumented agents feed the data base on a schedule or on demand. This subsection deals with information requirements for the Unit 2 ESD (SOA Monitoring).

3 Information for SOA Monitoring

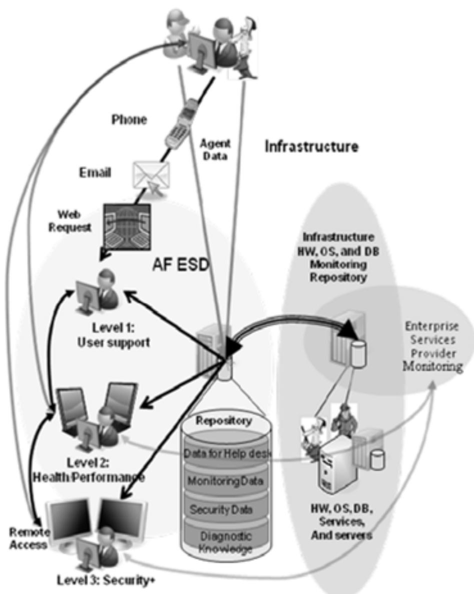


Fig. 2. Support Desk Operations

The knowledge base is where all information related to the enterprise SOA is stored. This will include the following:

- Hardware / software current status from DECC or APC
- Current reports on test activities including response times, frequency of test, etc.
- Current reports of usage data from service agent monitors and service logs, including number of users, session times, response times, etc.
- Hardware / software historical data
- A list of current alerts for the entire enterprise
- Historical data on alerts.

4 Customer Support and Help Desk

After initial self help is unsuccessful, the client will turn to the helpdesk. There are three levels of service provided. Level 1 will be the entry point to the system. At this level a client will call the helpdesk about a problem or other issue. The helpdesk assistant will collect some information, such as name, address, rank, SSN, etc. from the client, verify the client's identity, and attempt to quickly solve the issue using a pre-defined script and access to the KR. If this does not resolve the issue, it is categorized and escalated to the appropriate level (level 2). At level 2 a knowledgeable person will go through a more detailed script to collect information from the client and use the agent based monitoring process to gain more detailed information about the situation. Level 2 personnel will have more access and privileges, allowing them to make changes and attempt to fix the problem. If Level 2 personnel are able to solve the problem, they will close the TR and update the KR with information about the TR resolution. If the issue cannot be resolved by this level, the TR will be escalated to level 3, where in-depth analysis and coding are involved. If this resolves the issue, any code fixes will be pushed to the patch/Configuration Management (CM) team, and a temporary quick fix will be generated for future issues and added to the KR. This describes the basic helpdesk flow in the figures below.

The rest of the section is structured as follows:

- Levels of service
- Identity Verification System;
- Trouble Report Management System;
- KR System;
- Error reporting;
- Administrator duties, and
- Trouble report scenarios and data requirements.

Details of the agent-based architecture are complex and discussed in a separate paper [20].

5 Levels of Service

As previously described there are four basic levels of support, and three basic units or groups to assist in that support.

5.1 Level 0: Client Self-Help

Before executing the well-structured Help Desk Levels, there is a loosely defined "Level 0," which refers to the client's ability to solve their own issues through the KR or other means. The KR is primarily a helpdesk resource. By developing a culture within the client base of first checking the KR when an issue is encountered, this can further act as a filter, reducing the load on Level 1. The relationship of levels is shown in the next three figures

5.2 Level 1: Basic Information

Level 1 is where most of the TRs are created. All helpdesk requests will be recorded in a help desk TR. The Level 1 operator will follow a simple script, which will involve collecting basic information from the client, entering it into a TR, and creating the TR. At this point, the operator will start to address the client's issue. If the issue is resolved, the method of resolution will be entered in the TR. Often this will involve listing a link to the KR article that helped resolve the issue. Other times, it will be an operator text entry based on information personally provided to the client.

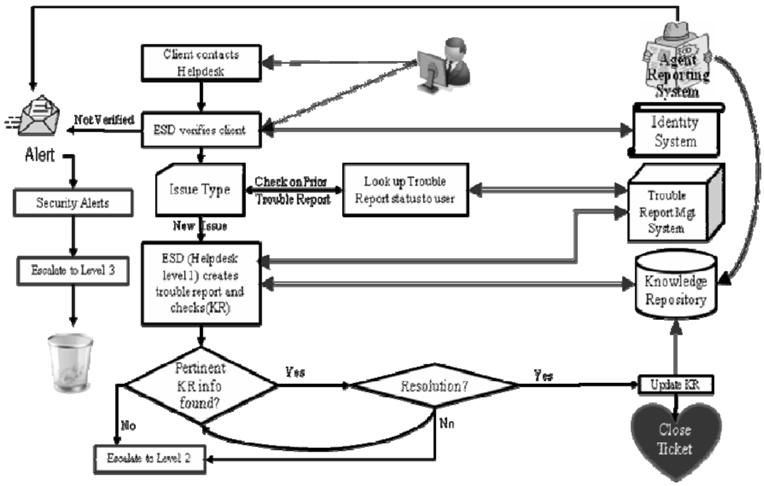


Fig. 3. Help Desk Level 1

This process ensures that all interactions are logged appropriately. An issue that is resolved quickly by a link to the KR is important to document. This provides positive feedback to the higher levels that create and maintain these KR articles. A high frequency of access can indicate that this is an important issue that clients need to understand. As a result, a more client-friendly KR article might be created to reduce call volume to the helpdesk or an enterprise-wide memo might be released addressing the issue. If the article addresses a bug and its workaround, a high access rate might indicate that this is a high priority for a permanent fix. In any case, the proper documentation of client requests, even for seemingly trivial issues, is an important part of the role of Level 1 operators.

In addition to the KR, Level 1 has access to basic live feeds from Units 2 and 3 of the ESD. These feeds indicate the current health, performance, and security situation of the network, and also include hardware and OS failure information. This feed is one-way, providing current information to the Level 1 operators. This information is intended for use when there are significant network issues. If a client calls and asks about a slow connection, this feed could provide a quick answer for that client if there is significant network traffic at the time. The TR could then be closed without escalation by verifying that Unit 2 knows about the issue and is working to fix it.

Not all issues will be resolved by a KR article link at Level 1. For more difficult problems and problems that require an administrator or other access, the TR is escalated to Level 2. Responsibility of the TR is passed to someone in Level 2, and Level 1's obligations for that TR are complete.

5.3 Level 2: Interactive Support

When a TR is escalated to Level 2, it is first scanned to determine the general issue. Although all operators in Level 2 are trained and capable of resolving issues, there will be some areas of expertise that people or levels develop within Level 2, so matching the problem with the person or group addressing it will be helpful for certain types of TRs. When someone has taken ownership of a TR, they contact the client by phone to authenticate them, confirm the information collected by Level 1, and collect more detailed information to address the issue. This interaction follows a script much like Level 1, but it is more detailed in terms of the authentication steps and the information collected. Level 2 operators all have the ability to remotely access a client's desktop. This is provided by agents running on the clients' desktops, which is part of the AF standard desktop configuration. If needed, the Level 2 operator will access the client's desktop, collect information directly from their desktop, and make changes to the client's desktop. Examples of changes would be software updates, installation of certificates or other security-relevant information, or removal of unauthorized device drivers or other software. All authorized software will be signed and certificated, so detection of unauthorized code should be easily accomplished by an automated scan.

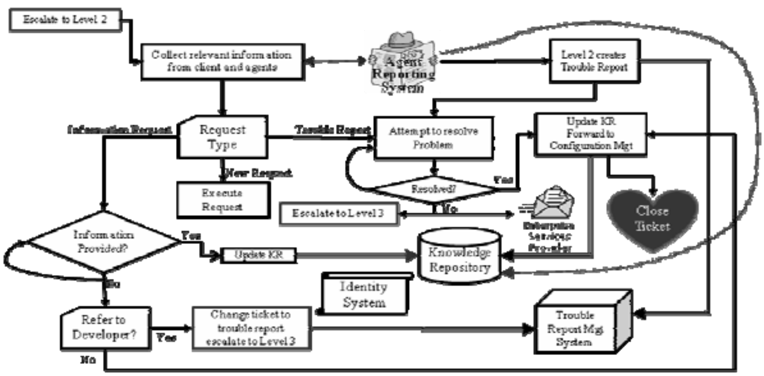


Fig. 4. Help Desk Level 2

The level of expertise at Level 2 is very high. The baseline training for Level 2 operators provides them an ability to resolve almost all issues faced. Specialized training for certain specific issues may be limited to a smaller group within Level 2, but this group will be consulted or given the TR when appropriate. Such specialized training might be for key services, such as the STS, MDE, or AD. Although the issues addressed might be difficult to diagnose and fix, by having these specialized groups within Level 2 for the most common problem area it is possible to avoid the very costly escalation to Level 3. However, sometimes there is just a fundamental problem with the code, either enterprise code or vendor code, or there is a fundamental architectural problem that must be changed, perhaps at the service provider facility not accessible directly from Level 2. In this case, the issue is escalated to Level 3 and/or sent to the service provider.

5.4 Level 3: Security, Serious Bugs and Vendor Support

Level 3 is not as well-defined as Levels 1 and 2. It consists of a collection of developers, security experts, architects, and vendor support channels for all COTS software and hardware. These experts are typically very costly to utilize, but their expertise and skills can resolve any issue, if properly diagnosed and allocated to them. The goal of the ESD is to address all TRs at the lowest level possible. Essentially, Level 1 acts as a filter to weed out all issues that are a waste of Level 2 resources. Level 2 acts as a further filter to weed out all issues that are a waste of Level 3 resources.

Level 3 does not actually take ownership of any helpdesk TRs. The one exception is security TRs opened by Level 3. The level 3 administrator, in consultation with other level 3 SMEs, may decide to call a Computer Security Incident Response Team (CSIRT) and is responsible for managing this team in its work. Level 3 is contacted by Level 2 when there is a serious issue to resolve, but Level 2 maintains ownership of the TR. Level 3 is treated more like consultants that the Level 2 operator can call on to solve serious problems. Although the TR does not change hands, the Level 2 operator contacts the client before escalating to Level 3 to inform them that their issue is more serious and may take some time to resolve. Level 3 issues are expected to take days or weeks to resolve instead of the minutes or hours expected at Level 1 or Level 2.

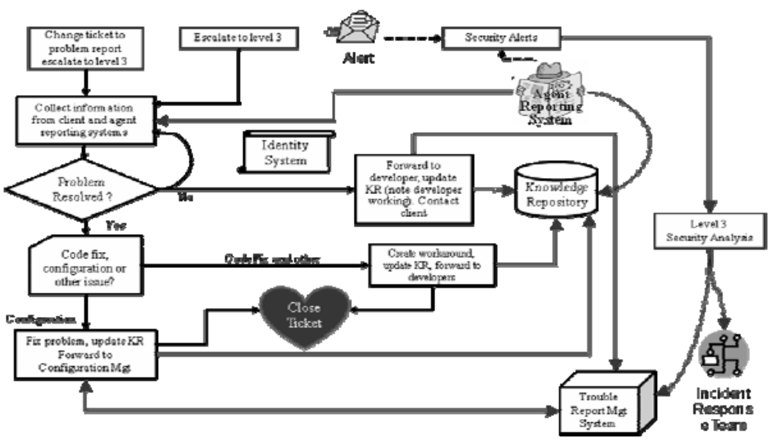


Fig. 5. Help Desk Level 3

6 The Knowledge Repository (KR)

The KR is a single integrated source of all information on the operation of the enterprise. It will be updated by all three units within the ESD and accessible to all three units within the ESD. The KR consists of a data base, and tools to import and categorize agent based data, as well as tools for search and discovery. The KR is access controlled at all levels.

6.1 Information for Help Desk Operations

The KR is where all information related to TRs is stored. This will include the following:

- Hardware/software current status from the service provider
- Hardware/software historical data
- A list of current alerts for the entire enterprise
- Historical data on alerts
- Reference material for all hardware components
- Reference material for all software components
- Reference material for all protocols and standards
- Configuration information
- All closed TRs
- All reference material that closed TRs link to
- Articles useful for setting up HW/SW, troubleshooting, etc.
- Frequently Asked Questions (FAQs)

Configuration information will consists of the standard desktop configuration, service configuration, infrastructure component configuration, and any other hardware, software, protocol, or standards settings or choices that have been established across the enterprise.

Closed TRs will be available in the KR. This will allow the helpdesk to search for similar TRs and use past experience to aid with current issues. Every TR will be added to the KR when it is closed. Access to these closed TRs will be discussed, since this is a potentially sensitive area. However, for now, it is sufficient to note that all closed TRs will be in the KR and available to someone, if not everyone.

In addition to closed TRs, any other documents referenced by the TRs should be included in the KR. Articles written internally by the ESD or other entities within the enterprise will be added to the KR if they would be useful for ESD activity. FAQs will be established and Configuration information will consists of the standard desktop configuration, service configuration, infrastructure component configuration, and any other hardware, software, protocol, or standards settings or choices that have been established across the enterprise. This includes the choice of which ports to open or close, when to use the various WS-* protocols, and which services should run at startup on a Windows Vista machine. For hardware, software, protocol, and standard elements there are an associated set of configurations and settings.

Closed TRs will be available in the KR. This will allow the helpdesk to search for similar TRs and use past experience to aid with current issues. Every TR will be added to the KR when it is closed. Access to these closed TRs will be discussed, since this is a potentially sensitive area.

In addition to closed TRs, any other documents referenced by the TRs should be included in the KR. This should be done not when the TR is closed but when the links are added to the open TR. The data should be added to the KR, then the link should be established, and then when the TR is closed the data and link are already there, so no additional work needs to be done. External links (web sites, etc.) will be captured in some way (like a Google cache) so that access is assured even when network connectivity is limited.

Articles written internally by the ESD or other entities within the enterprise will be added to the KR if they would be useful for ESD activity. FAQs will be established and probably be used when searching for general information on a topic, which might include KR data as well as other information. The KR search will be more for clients contemplating starting a TR or ESD/ESU personnel who want specific information for an existing TR.

Different levels of access will be established. These would correspond to different groups and roles that determine KR access. For TRs, the following levels are defined:

1. No closed TRs are viewable by callers
2. Callers can only view their own TRs
3. All TRs without any sensitive information are visible.
4. All TRs are visible (exception – security sensitive TRs), but fields designated sensitive blanked out
5. Everything is visible

The ability to update the KR will be restricted to level 2 and level 3 personnel. Level 1 is essentially a firewall for these people, weeding out all the previously solved and documented issues. Because of this, level 1 will only read the KR, much like the enterprise clients. When a new issue is found that cannot be addressed by the existing KR information, it is resolved at level 2 or level 3. After this issue is resolved, the level 2 administrator writes enough information in either the TR or a KR article so that the issue can be discovered in the KR and resolved by level 1 support the next time it happens.

For example, if a specific configuration must be used to access a particular new service, level 2 will field all calls relating to this issue until a KR article is added or the prior TRs provide enough information for level 1 to solve the problems. As new issues are discovered, the level 2 people will continue to resolve them and add more information to the KR. Eventually, a client attempting to access the service will find a wealth of information either through reading articles specifically about configuration for the service or through reviewing closed TRs with similar problems.

The main goal is to solve problems as few times as possible, since every level an issue is escalated is more costly than the previous one. The most preferable is for clients to solve their own problems through KR articles. When this is not possible, level 1 tries to resolve it. If it is a new issue, level 2 must address it. If it is a fundamental change, level 3 will be involved. However, once an issue makes it to a certain level, it should stop at a lower level the next time it arises, since the expert attention provided last time should be documented in a way that the levels below can use it without escalating the issue again.

Although only level 2 and level 3 can update the KR, anyone can suggest something be added to the KR. A simple way to implement this is through a web page, where any client can enter the information for a KR article, submit it to the appropriate group at level 2, and let the people at level 2 review the article and take the appropriate action. This allows people to provide input on issues they may have that do not get addressed by the KR, perhaps because they are low frequency issues or relate to a specific service that a client not in the ESD happens to know very well. The final say is by the level 2 person, though, since they ultimately are the maintainers of the KR.

7 Summary

This paper has presented a multi-tiered approach to enterprise support services. To our knowledge, it is the first full enterprise service support desk that is integrated with the hardware and operating systems of the support provider. These techniques are currently being implemented by DISA and the USAF in the DISA DECC stand-up of the enterprise solution for the USAF.

References

1. Air Force Information Assurance Strategy Team, Air Force Information Assurance Enterprise Architecture, Version 1.25, SAF/XC, April 11 (2008)
2. AFPD 33-3 Information Management, AF Portal Community of Practice: AF Information & Data Management Strategy – Implementation (Policy), `http://www.e-publishing.af.mil/`
3. COI Primer, AF Portal Community of Practice: AF Information & Data Management Strategy – Implementation (COI Primer)
4. DoD Directive 8320.2 "Data Sharing in a Net-Centric Department of Defense" and DOD Guidance 8320.2-G "Guidance for Implementing Net-Centric Data Sharing", AF Portal Community of Practice: AF Information & Data Management Strategy – Implementation (Policy)
5. Metadata Concept, AF Portal Community of Practice: AF Information & Data Management Strategy – Implementation (Metadata)
6. Transparency Integrated Product Team (TIPT) information and proceedings AF Portal Community of Practice
7. AFMAN 33-223, Identification and Authentication
8. AFMC Supplement 1, AFMAN 33-223, Identification and Authentication
9. CJCSI 6212.01D, Interoperability and Supportability of Information Technology and National Security Systems
10. DoDD 5000.1, The Defense Acquisition System
11. DoDD 4630.5, Interoperability and Supportability of Information Technology and National Security Systems
12. DoDD 8000.1, Management of DoD Information Resources and Information Technology
13. DoDI 8115.02, Information Technology Portfolio Management Implementation, October 30 (2006)
14. Joint Concept of Operations for Global Information Grid NetOps, Ver. 3, August 4 (2006)
15. Guide to Secure Web Services: Recommendations of the National Institute of Standards and Technology, NIST-US Department of Commerce Publication (August 2007)
16. Middleton, I.: Key Factors in HelpDesk Success (An analysis of areas critical to helpdesk development and functionality. British Library R&D Report 6247 (1996)
17. Service management solutions White paper Deliver service excellence through the unique advantages of IBM Service Management solutions (September 2007)
18. Farrell, J.A., Kreger, H.: Web services management approaches. IBM Systems Journal 41(2) (2002)
19. Managing Information Access to an Enterprise, Information System Using J2EE and Services, Oriented Architecture, IBM Redbook (January 2005)
20. Simpson, W.R., Chandersekaran: CCCT2010. An Agent Based Monitoring System for Web Services, Orlando, FL, April 2011, vol. II, pp. 84–89 (April 2011)
21. Introduction to Amber Point SOA Management System, Software Manual (December 2007)

A Study of Rapid Business Application Development in the Cloud

Jie Cui[1], Jing Min Xu[1], Huiping Lin[2], Weiping Li[2], and Zi Mu Sun[2]

[1] IBM Research – China, Beijing, China
[2] School of Software and Microelectronics, Peking University, Beijing, P.R.China
cuijie@cn.ibm.com, xujingm@cn.ibm.com,
linhp@ss.pku.edu.cn, wpli@ss.pku.edu.cn, zmsun@ss.pku.edu.cn

Abstract. The emergence of Cloud Computing promotes a rapid business application development (RBAD in short) paradigm, which enables non technical developers to create applications using the interactive and feature-rich tools via browser without caring about the underlying infrastructure. In this paper, we review the RBAD platforms in marketplace with various dimensions. To better understand their feasibility compared to the traditional software development, we conducted a user study with 12 participants creating application on 3 selected RBAD platforms. Analyzing their behaviors and feedback from different development perspectives, we design a conceptual business application model for RBAD platform, identify the best practice and gaps, and propose 6 design recommendations for RBAD system.

Keywords: Cloud, Business Application Development, User Study, PaaS.

1 Introduction

Cloud Computing dramatically reduces the complexity of owning infrastructure, software and services through on demand resources provisioning over virtualization and pay-as-you-go usage pattern [1, 2]. It also promotes a new paradigm in which business application is developed online and can be conducted by so called business developers who know the business knowledge best but with limited programming skills. Some serious business developers may possibly write simple script for enhanced business logic, just like they did in editing formula upon spreadsheet or defining email-filters. This new development approach greatly takes advantages of Cloud Computing and Software as a Service (SaaS) model in which vendors host and provide access to a software application over Internet [18]. Different from the traditional enterprises application development and deployment with a dedicated IT architecture and costly resource allocation, developers can build business applications faster and easier than ever without installing any tools, caring about the environment maintenance and compromising scalability, performance and security.

As the IT infrastructure gets increasing complex and new development techniques are emerging, this paradigm is becoming more and more attractive to small to middle size company as well as departments within larger organizations, because it lowers

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 398–407, 2011.
© Springer-Verlag Berlin Heidelberg 2011

the barrier of creating applications and can greatly reduce the development complexity and cost, without new investment on capitals and human resources.

There are many such rapid business application development (RBAD in short) platforms in market ranging from the ones which can step you through a drag-and-drop interface to build form-based application such as Wufoo [15], JotForm [18], FormAssembly [12], Zoho Creator [14], Quickbase [19] etc, to sophisticated platforms like Force.com [6] (a product of Salesforce.com), WaveMaker [16] and Bungee Connect [20], which can integrate a company's existing programs through Web interfaces and provide libraries, scripting or server side programming language for business process customization. As more big companies such as Google, Microsoft, IBM and Amazon join this fray, fostering the rapid market growth, we envision this new paradigm as a potential shift for business application development and corporate IT and might be disruptive to the traditional software development.

To investigate the feasibility of this programming and business model and also analyze requirements and feedback from real user experience, we overview the emerging marketplace with a variety of dimensions. In particular, a user study is designed and conducted on the selected representative platforms: Zoho Creator, Force.com and WaveMaker. Furthermore, we explore and discuss the following research question: What are the associated implications and design recommendations for RBAD system?

The structure of the paper is as follows. Section 2 presents the related work. A review of the RBAD platforms in market is given in Section 3. Section 4 details our user study design and conduction, and is followed by the result analysis in Section 5. Section 6 further discusses the research questions with critical insights gained. Finally, we summarize the paper and outlook the future work in Section 7.

2 Related Work

Gautam Sharoff named this new paradigm as Dev 2.0 [3], where the line between users and developers is blurred, and an application is available to all stakehoders through the lifecycle as it evolves. More generally, it is referred as Platform as a Service (PaaS) model which takes SaaS model one step further, providing end to end application development environment from editing code to debugging, deployment, runtime and management [4, 5]. PaaS model is built upon Cloud infrastructure and uses a multi-tenant deployment and development tools providing important value for developers to focus on development and innovation, while eliminating the burden of configuring servers, implementing management tools, wresting with the storage, OS and middlewares, and dealing with the network interfaces. Force.com is an industry leading PaaS offering. It further categorizes PaaS solutions into 4 categories [23]: 1) Social application platforms (e.g., Facebook [24]) which provide APIs so third parties can write new application functionality that is made available to all users. 2) Raw compute platforms like Amazon Web Services [21] which provide storage, processor, and bandwidth as a service. 3) Web application platforms like Google App Engine [22] which provide APIs and functionality for developers to build Web applications. 4) Business application platforms like Force.com which provide application infrastructure specifically geared toward transactional business applications.

In this paper, we are more interested in the last PaaS type used by business developer instead of professional developers who are mostly targeted by the other three types. Although there are quite a lot researches in investigating the issues of developing software in a Cloud environment or the design factors of PaaS platforms [25, 26, 27], to our best knowledge, there has been relatively little research giving systematic survey and exploratory study on RBAD in the Cloud to discuss the feasibility and real user feedback. Therefore, our paper can complement to this research field.

3 Overview of the RBAD Platform

We totally studied 9 RBAD platforms and reviewed them from a variety of aspects as follows: 1) target user, 2) target application, 3) functions, 4) developing lifecycle, 5) differences from local IDE (Integrated Development Environment) .

According to our survey, the RBAD platforms can be divided into two categories, i.e., 1) online form builder with or without database (DB) support, and 2) sophisticated platforms supporting customized applications with more complicated business processes.

3.1 Online Form Builder

Two types of online form builders were identified depending on whether they have the support of DB operation. Totally 7 online form builders were selected and studied.

1. Online Form Builder without DB Support

Online form builder without DB support is more or less an online service that helps users create forms without writing codes and provides an interface for collecting data, but little DB operations are allowed, e.g., Wufoo [15], JotForm [18], FormAssembly [12], and Iceberg[28]. Their target users are those that have little programming skills and little time to learn.

2. Online Form Builder with DB support

Online form builders with DB support allow users to build online database application with CRUD (create, retrieve, update and delete) controllers that can deal with more complicated data. Three platforms, namely QuickBase [19], DabbleDB [13], and Zoho Creator [14] were studied. Comparing with online form builders without DB support, their target applications usually include more complicated data and report analysis. They provide more functions to users such as creating the database, sharing the information and carrying out report analysis.

3.2 Sophisticated RBAD Platform

Two more powerful platforms Force.com [6] and WaveMaker [16] were studied as well to see how it can support customized business logic and presentation.

Force.com platform is a PaaS product owned by Salesforce.com, who delivers CRM solutions to customers over internet as service. The platform allows external developers to create add-on applications that integrate into the main Salesforce application and are hosted on Salesforce.com's infrastructure. Force.com targets at both business developer and technical developers.

WaveMaker is available as a traditional product that is installed on a developer's computer (desktop) or as a software as a service that is accessible via the web (cloud). The difference with other platforms is that it can import local database and automatically generate data model you want to access on WaveMaker.

4 User Study

We designed a typical online airline reservation application scenario and conducted a user study for developing it on the representative RBAD platforms to identify advantages and functional gaps comparing to traditional application development from real user feedback.

4.1 Online Airline Reservation Application

The application scenario is quite straight forward and can be simply described as the following steps.

1. User is allowed to search for the flight and airfare by selecting origin, destination, dates and flight type.
2. User can select an available one and create an order.
3. After assigning the passengers information and inputting the payment information, user can submit the order.
4. The credit card will be validated and then a notification email will be automatically generated and sent to the user.

We choose this application for testing the platform functions not only because it represents the typical features in reservation systems including search, list and order, but also because it covers most of the test points we designed for RBAD platforms as follows.

1. The diversity of the field types in business objects. E.g., the flight information contains field type of date, text, currency and integer etc.
2. The correlation of business objects. E.g., an order instance associating flight information and probably user information would be created after user submits the order.
3. Navigation logic in the application. This application includes a typical sequential page flow from the starting search page to the final order page, in which parameters are passed to ensure the page linkage and integrity. In traditional Web application development such as using JSF (Java Server Face), page flow can be explicitly configured in an XML file which indicates the “from” and “to” pages and associated navigation conditions. We will see how it can be achieved in RBAD platforms.
4. The ability to integrate external application. We intentionally provided two external Web Services based on SOAP and REST respectively, one to verify user credit card, and the other to send email notification.

4.2 Participants and Procedures

We recruited 12 volunteers (10 male, 2 female) of college students from different majors (e.g., business, mathematics, and computer science (CS)). The reason that we also hired the students with programming background is to testify the advanced functions that the platform provides such as programming for customized logic. 4 students whose major is CS experienced with software development project for 2 to 4 years. Other students didn't have much Web programming experience.

We balanced the programming experience across the three groups, see table 1.

Table 1. User study group assignment

	Zoho Creator Group	Force.com Group	WaveMaker Group
Participants	4	4	4
CS students	1	1	2

For each group, there was a facilitator assigned to manage the whole procedure and observe their behaviors. Firstly an overview of the tasks they would be carrying out was given, followed by a 30 minutes tutorial, a more concise version of the quick start guide of each platform, with a hands-on session to get them familiar with the functions of the platform. Then a questionnaire adapted from [7] to the specific task of RBAD was completed by participants to identify their expectations of the outcomes and emotional reactions to the unfamiliar software.

A requirement documents describing the application scenario (Section 4.1) was provided to the participants. Also, an additional tutorial would be given when participates made no progress. After participants confirmed they completed the task, a focus group discussion was conducted to gather their feedback, and a survey questionnaire designed to identify the issues from different development perspectives was finally completed.

4.3 Environment

Zoho creator and Force.com platform provide free version and 30 days free trial version respectively with almost all the functions claimed except advanced performance and security insurance. Also, free registered user can directly access WaveMaker cloud version. Therefore, a shared account was applied on above platforms by each group. Additionally, each group was separately located and was allowed to discuss the application design, including business objects construct, business logic, UI design and page flow, which was guided by the CS background students. All the participants were required to develop as least one component of the application and the procedure was recorded. The connection speed to internet was around 1M but sometimes it was deceased due to network traffic.

5 Result Analysis

We observed how participants approached the task and classified their development activities into 4 types 1) data (setting up data model on Zoho Creator and Force.com or connecting local DB on WaveMaker, viewing or updating db records and etc) 2) business logic (defining or testing business objects association, workflow and app integration) 3) UI (pages and page flow creation, editing, viewing or testing) 4) problem solving (referring back to the tutorial or online help, asking for partner's or facilitator's help, and debugging a error encountered). Since the facilitators with strong CS background had a pre in-depth study of the platform, they can be quite easy and accurate to detect the distinctions of the above behaviors.

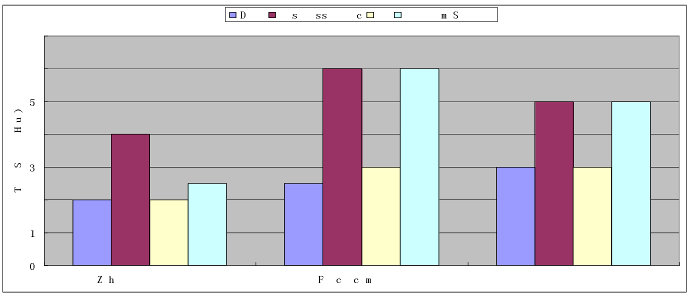


Fig. 1. Development time spent in different activities

Figure 1 shows the time spent in those four types of behaviors of each group. Basically, user spent most of time in developing business logic and UIs, that's also very common in traditional development paradigm and specially in this experiment which evolved scripting (e.g., Deluge on Zoho Creator) or additional programming (e.g., Apex [8] and Visualforce [9] on Force.com, JavaScript and Java on WaveMaker). Creating data structures seem not time consuming, suggesting the facilitation for building forms is very effective, e.g., drag-and-drop tool (Zoho Creator), point-and-click tool (Force.com) and local DB connection (WaveMaker). Additionally, we noticed users on Force.com spent significantly more time on problem solving than the other two so that we believe there is a relative steep learning curve for novice users on Force.com. That is also reflected in a large amount of documentation on developerforce [10] that details the advanced functionalities.

5.1 Data Development

Data model is the basic construct of an application especially for the business application which mainly manipulates information entities that transform in the business to achieve the organization goal. Since the majority of the audience are business developers or non-technical administrators, many RBAD platforms use terminology that is familiar to them, e.g, it is avoided to use "DB Tables" and "Rows" instead of "Objects" and "Records" within the context of Force.com; Application creation is started from “Form”, “Fields” and “View” which are more visualized from the end user view on Zoho Creator. Compared to the traditional DB development, Table 2 summarizes the function supportability of the 3 platforms.

Table 2. Data development supportability

	Zoho Creator	Force.com	WaveMaker
Related objects	supported	supported	supported
Graphic widgets	supported	supported	supported
Data validation	supported	supported	supported
CRUD operation	supported	supported	supported
Data import	XLS,CSV, MDB	CSV	Local DB

5.2 Business Logic Development

The powerfulness of the supporting business logic layer is a good demonstration point in terms of platform flexibility. The business logic orchestrates interaction among business entities, therefore controls the navigation of user interface. The server side scripting and programming capabilities allow advanced user to develop relative complex and customized business logic on the 3 platforms. We further categorized the business logic support in Table 3.

Table 3. Business logic development supportability

	Zoho Creator	Force.com	WaveMaker
Form actions	supported	supported	supported
Workflow editor	not supported	supported	not supported
App integration	partially supported	supported	supported
Programming	Delug	Apex, VisualForce	Javascript Java Service

Regarding to application integration, Zoho Creator provides REST as well RPC style Data APIs that allows external websites to collaborate with its applications, but with limited support for calling external REST or Web Service. Force.com has the two-way integration approach which means the external application can access Force.com data and logic through Web Service API, and more facilitation such as data validation, workflow, Apex triggers and the Force.com security scheme can help user to integrate external application within Force.com application. While WaveMaker provides the user interface for importing REST, Web and RSS services to the application which is more convenient than writing code.

5.3 UI Development

RBAD platforms provide metadata-driven, automatically generated user interfaces and page layouts which impress the participants most. They found it was fun to drag and drop visual widgets and compose or layout them in the canvas. More advanced functions are summarized in Table 4.

Table 4. UI development supportability

	Zoho Creator	Force.com	WaveMaker
Drag & drop	supported	supported	supported
Custom theme and layout	supported	supported	supported
JavaScript	not supported	supported	supported
CSS	not supported	supported	supported
Page flow configuration	not supported	not Supported	not Supported

Visualforce [9], a component-based framework in MVC paradigm provided by Force.com, helps users to create and deploy fully custom user interface. Not like fully customized UI on Force.com, the other two platforms provide most necessary UI customization capabilities such as theme or layout selection, or user defined CSS (WaveMaker) etc.

5.4 Summary

Analyzing the comments from the participants, we summarize the 3 platforms as below.

1. **Zoho Creator is the most intuitive platform on which participants saved half of the time than the other two.** Programming script was also an enjoyable task for participants and could be quickly referenced in the documentation, but not being able to call external Web Service is an imperfection from application integration view.
2. **Force.com is the most powerful and comprehensive platform among them.** All the functional requirements of the airline reservation can be implemented on Force.com including form creation, Web Service calling and page flow. But it took participants much longer time to build the application and they complained about the long learning process and always got frustrated in facing errors without quickly identifying the problems and always lost in the numerous documents Force.com provides.
3. **WaveMaker targets more technical people who should have the knowledge of database and programming.** The performance was not being satisfied since every action was needed to manually saved for a long time and data was always lost which heavily decreased the development productivity.

6 Conclusion and Future Work

In this paper, we have presented an overview of the new and promising development paradigm namely RBAD and a user study for the representative RBAD platforms. Through the analysis of user behavior and feedback, we contributed to the evidence and identification of both advantage and gaps over traditional development. We also distilled 6 design recommendations for future RBAD and PaaS platform. We plan to further develop the work in three directions. Firstly, we will carry out additional empirical work to explore other functions such as security, performance, testing and deployment with larger set of RBAD platforms. Secondly, we will begin to develop lightweight tools to validate the business application model proposed and apply the design recommendations. Finally, we will expand the study to the PaaS domain and larger development community to identify problems and solutions

Acknowledgments. We sincerely thank the participants for volunteering their time to our user study.

References

1. Armbrust, M., Fox, R.G.A., et al.: Above the clouds: A berkeley view of cloud computing. Technical Report, UCB/EECS-2009-28, University of California, Berkeley (February 2009)
2. Clark, R.: A Break in the Clouds: Towards a Cloud Definition. ACM SIGCOMM Computer Communication Review 39(1), 50–55 (2009)
3. Shroff, G.: Dev2.0: Model driven development in the Cloud. In: ACM SIGSOFT 2008, Altanta, Gerogiza, USA (2008)
4. Lawton, G.: Developing Software Online with PaaS technology. IEEE Journals Computer 41(6), 13–15 (2008)
5. Mitchell, D.: Defining Platform-As-A-Service, or PaaS, `http://blogs.bungeeconnect.com/2008/02/18/`
6. defining-platform-as-aservice-or-paas/
7. Force.com, `http://www.salesforce.com/platform/`
8. Compeau, D.R., Higgins, C.A.: Computer self-efficacy: Development of a measure and initial test. MIS Quarterly, pp. 189–211 (June)
9. Salesforce.com. An Introduction to Force.com Apex Code., `http://wiki.developerforce.com/index.php/An_Introduction_to_Apex`; Salesforce.com. Introduction to Visualforce, `http://wiki.developerforce.com/index.php/An_Introduction_to_Visualforce`
10. DeveloperForce, `http://developer.force.com/`
11. Weissman, C.D., Bobrowski, S.: The design of the force.com multitenant internet application development platform. In: Proceedings of the 35th SIGMOD international conference on Management of data, SIGMOD 2009 (2009)
12. FormAssembly, `http://www.formassembly.com/`
13. DabbleDB, `http://www.dabbledb.com/`
14. ZoHo Creator, `http://creator.zoho.com`

15. Wufoo, http://wufoo.com/
16. WaveMaker, http://www.wavemaker.com
17. Chou, D.C., Chou, A.Y.: Analysis of a new information systems outsourcing practice: software-as-a-service business model. International Journal of Information Systems and Change Management 2(4) (2007)
18. JotForm, http://www.jotform.com
19. QuickBase, http://quickbase.intuit.com/
20. Bungee Connect, http://www.bungeeconnect.com/index.html
21. Amazon Web Services, http://aws.amazon.com/
22. Google. What Is Google App Engine (2009), http://code.google.com/appengine/docs/whatisgoogleappengine.html
23. Salesforce.com. Types of PaaS solutions, http://www.salesforce.com/paas/paas-solutions/
24. Facebook, http://www.facebook.com
25. Sledziewski, K., Bordbar, B., Anane, R.: A DSL-Based Approach to Software Development and Deployment on Cloud. In: 24th IEEE International Conferences on Advanced Information Networking and Applications (AINA 2010), Perth, Australia (2010)
26. Boniface, M., Nasser, B., Papay, J., et al.: Platform-as-a-Service Architecture for Real-Time Quality of Service Management in Clouds. In: 2010 Fifth International Conference on Internet and Web Applications and Services, Barcelona, Spain (2010)
27. Iceberg, http://www.geticeberg.com/

Design Innovation for Enterprise Software

Mark Detweiler and Liam Friedland

User Experience, Informatica Corporation, 100 Cardinal Way,
Redwood City, CA 64063 USA
{mdetweiler,lfriedland}@informatica.com

Abstract. This paper offers examples of how one User Experience (Ux) team has supported corporate innovations in enterprise software using a variety of design-innovation practices. It describes tactical and strategic design and usability activities that have been used to drive a variety of innovation activities - including incrementally refining existing products, unifying the look and feel of acquired products, and exploring and visualizing new products for existing and new users.

Keywords: Design, innovation, user experience (Ux), user interfaces, enterprise software.

1 Introduction

The purpose of this paper is to share insights about how one Ux team has supported innovations in enterprise software. Although the innovation literature is quite vast, and covers a wide range of different types of innovation [1] this paper is primarily focused on ways a Ux team can support product innovation as an effective strategy for adapting to changing technologies, markets, and competition. It should be noted that innovation in enterprise software frequently includes innovating around business and work processes; and routinely addresses issues related to governance, and the kinds of work practices and workflows that are allowed. As new enterprise functionality is envisioned and designed, it often involves changes in the way work is done, how people collaborate to do that work, the types of artifacts created, and the actions different people and roles can perform with or on these artifacts.

Rettig [2] and Sherman [3, 4] have cataloged some of the special challenges and issues associated with developing good user experiences for enterprise software. Two primary issues are the variety and heterogeneity of code that needs to interoperate, and the lack of involvement of real end-users in helping to specify what should be built. Lack of end-user involvement—especially involving "lead users" [5] may be the most challenging issue from a Ux perspective. Without end-user participation, Ux designers are forced to rely on second-hand information about users' needs from IT managers and procurement proxies. In addition to these, other challenges emerge related to the embedded base of technology, the different update cycles and migration considerations of enterprise technology, user skills and training support required to master new capabilities, and compatibilities with existing business processes and practices. In light of these considerations, it is easy to see why some enterprise-focused Ux teams feel they are fighting an uphill battle to deliver high-quality user experiences.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 408–414, 2011.
© Springer-Verlag Berlin Heidelberg 2011

2 Ux Organization and Infrastructure

A Ux team's ability to successfully adapt to constantly changing technology and business requirements is heavily influenced by how the team is organized, and the kinds of processes and infrastructure it develops and maintains. The Ux team at Informatica is distributed across six global locations, in four time zones. This distribution allows Ux designers to be co-located with development teams to facilitate close collaboration, communication, and coordination. However, it is critical that the Ux team maintain a shared vision, and use common tools to achieve that vision. Invariably the difference in time zones poses special challenges for holding regular team meetings, but such meetings are essential to keep everyone informed about the status of design and research activities, and to gather peer feedback in design reviews.

In addition to standard email and web-conferencing software, the Ux team relies heavily on a central document repository to share wireframes, design specifications, and all sorts of presentations and reports. The Ux team has also invested heavily in developing a UI pattern library, consisting of a formalized, web-accessible repository of visual and behavioral guidelines, examples, detailed UI widget descriptions, and recommendations regarding intended usage. The goal for creating this library of tools was to foster a common look and feel across products, facilitate Ux and other groups' communications, and to increase the chances of making users' experiences more consistent, particularly for people using multiple Informatica products.

Figure 1 illustrates a few of the most important of the many possible paths of innovation; for a similar characterization, see [6]. Existing and new markets /users are depicted on the y-axis and existing and new products and services are depicted on the x-axis. The bottom left quadrant is a common starting point for innovation. Innovation can proceed upwards towards the upper-left quadrant by trying to address existing markets / users by providing incremental enhancements to existing products and services, and trying to win access to and to grow new markets. Innovation can also proceed sideways towards the bottom-right quadrant, by trying to create new products and services that can be sold into an existing customer base. Innovation can also proceed diagonally towards the upper right quadrant by trying to develop new products and services for new markets / users (e.g., see [7, 8, 9, 10, 11]. The kind of innovation paths a company takes has a strong influence on the kinds of directions and activities a Ux team can take.

Our Ux design innovation processes have been shaped across both tactical and strategic initiatives: a) supporting incremental refinements to existing products and markets; b) unifying the look and feel of acquired products; and c) visualizing and developing new products for existing and new users and markets. Although each of these types of innovation has required tailored Ux activities, they all share a common goal of attempting to help our company create greater market share, promote a common look and feel, enhance usability, and increase customer satisfaction. In the next sections we provide concrete examples of what has worked well to promote these different types of innovation.

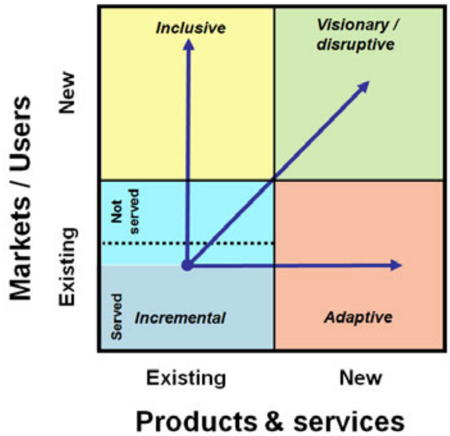
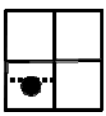


Fig. 1. Paths of innovation

3 Design and Enterprise Innovation

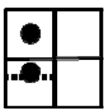
3.1 Extending Existing Products – Incremental Refinement



Having clear requirements, at the right level of detail, is critical to extending a product's functionality and incrementally creating quality user interfaces across releases. However, getting proper requirements at the right level of detail to influence design seldom occurs without a systematic process for including Ux concerns, (e.g., probing to understand users' current and future activities and goals, desired juxtapositions of information, and preferred workflows), as the requirements are gathered and analyzed. This realization led us to develop a standardized requirements process that we refer to as Requirements Analysis Process (RAP). RAPs provide a common requirements-gathering framework with flexibility to address particular needs of different product groups. RAPs are conducted at existing and potential customer sites, and remotely to understand stakeholders' roles, tasks, workflows, collaboration, pain points, and desires. In the best cases, RAPs have included opportunities to observe users at work on their own systems, with their own data. This type of observation has provided invaluable insights into users' work practices that are simply not possible to uncover with more traditional usability studies [12]. RAPs also provide an initial set of prioritized features and functionality that can be systematically compared across different stakeholders and customers [13]. Further, we have used artifacts gathered during RAPs to front-load an Agile-development process in which additional requirements and clarifications are gathered "just in time" to support Agile development sprints.

In addition, we have used iterative prototyping as the primary means of specifying new refinements [14, 15, 16, 17]. This prototyping has tended to be more tactical than strategic, in the sense that its primary purpose has been to create concrete visual representations of design concepts, to verify that customers' needs have been accurately captured, and to support development work across sprints for a single release. As the prototypes evolve, they are transformed, progressively, into design specifications that serve as a source of truth for how the coded user interfaces should look and behave. Development teams then use these specifications to guide their coding, and QA teams use them to see how well the developed code matches the intended look and feel. Finally, we have also used traditional in-lab and remote usability testing, (e.g., at alpha and beta stages of a product release) to identify opportunities to increase design quality and improve the user experience.

3.2 Unifying the "Look and Feel" of Acquired Products



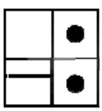
Over the past 5 years, Informatica has acquired 8 companies with products of varying complexity and maturity. While these acquired products and capabilities extended Informatica's core data-integration solutions into new markets and types of users (e.g., identifying and improving data quality, exchanging B2B data, managing lifecycle data, processing complex events, managing master data, and fostering ultra messaging), they presented the Ux team with unique challenges and opportunities.

In response to these challenges, the Ux team adopted some common tactical and strategic engagement practices. After establishing good working relationships with the product leadership team, the Ux team performs a Ux Audit to identify key strengths and opportunities for unifying and refining a product's visual and interaction design and usability. A Ux Audit is similar to a heuristic evaluation, but involves examining a number of additional factors often absent in heuristic evaluations, (e.g., assessing visual branding, identifying thick- and thin-client UI patterns that differ from those used in the company's other products, considering implications of conceptual and functional redundancy, assessing how well the user interface will scale, and probing issues related to maintainability, (e.g., managing roles and permissions from a common administration console)).

Typically, a Ux Audit begins with one or more demonstrations of product functionality, and discussions with Product Mangers to identify a list of key tasks and areas to explore in depth. Then two to three members of the Ux auditing team work through the listed functionality and tasks, interactively, over multiple working sessions. During these sessions, issues are first identified, and then later categorized and prioritized. The results of the Ux Audit are written up as a report, with lots of heavily annotated screenshots, and descriptions of the discovered issues. This report is normally sent to both the Product Management and Development organizations to review, and is followed up with a joint presentation of the issues, and a question-and-answer session.

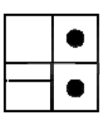
In most cases, Development teams do not have sufficient time or resources in their current release to address all the issues uncovered in a Ux Audit report. Therefore, the Ux team advises and supports a phased approach to acting on prioritized changes. Phase 1 often involves developing a new visual design and re-skinning the acquired product, (e.g., unifying logos, banners, color sets, splash screens, and desktop icons). This re-skinning reinforces the company's branding goals, and unifies the way the UI looks so that true family resemblance is achieved with the company's other products. Phase 2 often involves complementary design and research activities. Design activities focus on helping address issues that can be changed somewhat easily to increase consistency and to improve usability, some deeper visual re-skinning, and suggesting ways to convert some of the acquired product's existing UI widgets to easy-to-consume UI patterns. Research activities attempt to develop deeper understanding of users' goals, tasks, collaboration patterns, and workflows. This is also the time when user testing is conducted to provide richer insights and data to inform subsequent design and prototyping. Phase 3 requires more substantial efforts to integrate the acquired product's code base into either web- or Eclipse-based user-interface frameworks, while adopting and extending patterns in the Ux pattern libraries.

3.3 Visualizing Future Products, Markets, and Users



In contrast to the first two types of innovation, creating new products has entailed extending Ux tools and activities. First, we have found it extremely valuable to partner with subject matter experts (SME's) with deep domain knowledge. We often conduct deep dives with SMEs, and hold joint brainstorming sessions and workshops to transfer both domain and tool knowledge, as well as to inspire us to think about solving customers' problems in new ways. Second, we have profited from conducting competitive analyses for new products somewhat differently than for existing products, (e.g., thinking about markets and users beyond the traditional categories that shape current products and solutions). Third, we have invested time to spot and analyze technological, business, and social trends [18, 19], and then to socialize and debate learning from these explorations within and outside the Ux team.

3.4 Vision / "Concept Car" Prototypes



We recognized that other types of "vision" or "concept car" prototypes could be critical to help paint a longer-term vision for the user experience of one or more products. The rationale behind creating vision prototypes is similar to creating planning scenarios [20], (e.g., to explore a variety of possible futures that move beyond current concerns and problems). These more strategic prototypes are intended to provide longer-term perspectives and opportunities to think outside of the typical constraints imposed by limitations in existing technology and infrastructure, and schedule-driven releases. As such, they are meant to be future-oriented and to embody visual and behavioral possibilities that help evoke compelling representations and experiences of where multiple products might go. They help challenge current work and examine unstated assumptions about how user interfaces and user experience might unfold. Typically, vision prototypes involve creating stories and then fleshing out user experiences and UI approaches to bring these stories to life (cf. [21]). The stories or scenarios tend to begin as activities along a rough timeline, but then quickly evolve to show how various artifacts are created and transformed as different roles work together to achieve common goals. These are serious visual thinking tools meant to stimulate cross-disciplinary discussions, and ultimately to help set direction for future investments.

4 Conclusion

In this paper we have attempted to share examples of Ux activities that have helped support a variety of enterprise innovations. We recognize that each company will face unique challenges, based on its specific corporate culture, level of executive support for Ux staffing and resourcing, and general receptivity to design and usability throughout the organization. However, we hope these examples will inspire and help other Ux teams think about ways to not only face the complex tradeoffs that are an inevitable part of designing enterprise software, but come closer to addressing end users' real needs.

References

1. Keeley, L.: Innovate! Now (2009), http://www.innovatenow.us/wdkm_in/wcm/resources/documents/adobe_pdf/2009_innovation_summit/4237_2269518_22600/innovationsummit09_keeley.pdf
2. Rettig, C.: The trouble with enterprise software. MIT Sloan Management Review 49(1), 21–27 (2007)
3. Sherman, P.J.: The user experience of enterprise software matters (2008), http://www.uxmatters.com/mt/archives/2008/12/the-user-experience-of-enterprise-software-matters.php
4. Sherman, P.J.: The user experience of enterprise software matters, Part 2: Strategic user experience (2009), http://www.uxmatters.com/mt/archives/2009/03/the-user-experience-of-enterprise-software-matters-part-2-strategic-user-experience.php
5. von Hippel, E.: Lead users: A source of novel product concepts. Management Science 32(7), 791–805 (1986)

6. Jacoby, R., Rodriguez, D.: Innovation, growth, and getting to where you want to go. Design Management Review 18(1), 10–15 (2007)
7. Cash Jr., J.I., Earl, M.J., Morison, R.: Teaming up to crack innovation & enterprise integration. Harvard Business Review 86(11), 90–100 (2008)
8. Cheskin, Fitch: Worldwide: Fast, focused & fertile: The Innovation evolution (2003), `http://www.cheskin.com/cms/files/i/articles//3__Cheskin_Fitch_Innovation_rpt.pdf`
9. Christensen, C.M.: The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail. Harvard Business School Press, Boston (1997)
10. Anthony, S.D., Johnson, M.W., Sinfield, J.V., Altman, E.J.: The Innovator's Guide to Growth: Putting Disruptive Innovation to Work. Harvard Business Press, Boston (2008)
11. Rothaermel, F.T., Hess, A.M.: Innovation strategies combined. MIT Sloan Management Review 51(3), 13–15 (2010)
12. Fulton Suri, J., Howard, S.G.: Going deeper, seeing further: Enhancing ethnographic interpretations to reveal more meaningful opportunities for design. Journal of Advertising Research 46(3), 246–250 (2006)
13. Botta, R., Bahill, A.T.: A prioritization process. Engineering Management Journal 19(4), 20–27 (2007)
14. Coughlan, P., Suri, J.F., Canales, K.: Prototypes as (design) tools for behavioral and organizational change. The Journal of Applied Behavioral Science 43(1), 1–13 (2007)
15. Fulton Suri, J.: The experience evolution: Developments in design practice. The Design Journal 6(2), 39–48 (2003)
16. Houde, H.C.: What do prototypes prototype. In: Helander, M., Landauer, T., Prabhu, P. (eds.) Handbook of Human-Computer Interaction, pp. 367–381. Elsevier, New York (1997)
17. Lim, Y.-.K., Stolterman, E., Tenenberg, J.: The anatomy of prototypes: Prototypes as filters prototypes as manifestations of design ideas. ACM Transactions on Computer-Human Interaction 15(2), 1–27 (2008)
18. Ofek, E., Wathieu, L.: Trends that could shake up your business? Harvard Business Review, 125–131 (July-August 2010)
19. Kisker, H., Ried, S., Shey, H.: The state of enterprise software and emerging trends. Forrester report (2010)
20. Fahey, L., Randall, R.M. (eds.): Learning From the Future: Competitive Foresight Scenarios. John Wiley & Sons, Inc., New York (1998)
21. Moggridge, B.: Innovation through design. In: International Design & Culture Conference (2008)

Company Culture Audit to Improve Development Team's Collaboration, Communication, and Cooperation

Emilie Gould[1] and Aaron Marcus[2]

[1] SUNY Albany, Albany, NY
egould@albany.edu
[2] Aaron Marcus and Associates, Inc., 1196 Euclid Avenue,
Suite 1F, Berkeley, CA, 94708 USA
Aaron.Marcus@AMandA.com

Abstract. Multinational companies worldwide seek to improve product/service development-teams operating in different countries. One of their objectives is to improve collaboration, communication, and cooperation among teams. Aaron Marcus and Associates, Inc. (AM+A) recommended countries and locations in which to undertake research into teamwork within a company, accounting for cultural differences and approaches to teamwork/leadership

Keywords: Business, collaboration, communication, cooperation, culture, design, development, software, management, user interface, user experience.

1 Introduction

A California client planned a company-wide study of collaboration and asked AM+A to recommend four sites from a list of 15 countries in six work regions. The client had already chosen a specific site in the US because, as the client's corporate headquarters, it needed to be able to work with sites throughout the world. A key assumption of our recommendations was that the other sites should represent the *greatest diversity relative to the headquarters site in terms of culture*, along with the *greatest opportunities for innovation in collaboration.*

Initially, we used two sets of criteria to identify appropriate sites. The first set of criteria were based on traditional measures of culture that are applicable to teamwork. The second set dealt with technological innovation and globalization. After an informational interview with the project sponsors, additional criteria were discussed and additional data on the client employee demographics were made available. As a result, an additional criterion was added in our data summary: gender balance.

2 Recommendations

AM+A recommended that the four cities to be studied along with the US corporate headquarters site should be the following:

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 415–424, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Amsterdam	Netherlands	Representative of North European, gender-balanced, egalitarian styles of collaboration
Bangalore	India	Representative of both S. Asian mentor-based collaboration and new modes of globalization
Dubai	United Arab Emirates	Representative of high power distance and new modes associated with globalization
Tokyo	Japan	Representative of East Asian, Confucian-based collaboration styles (high power distance and strong work group orientation)

Alternative cities included:

- Feltham for Amsterdam: Because the US and UK share many cultural values, Amsterdam was chosen as representative of more egalitarian, gender-neutral styles.
- Singapore for Dubai: Dubai currently leads the world in population growth as the UAE moves from oil to a new economy based on trade, finance, and information. Singapore has already successfully refocused its economy on the same elements through a much more regulated and state-controlled process representative of traditional Chinese values (strong leadership, high power distance, and collectivism).
- Seoul for Tokyo: Like Tokyo, Seoul uses a strongly Confucian style of leadership that promotes group coordination and tends to limit individual action.

If circumstances allowed further cities, Shanghai could be researched. It was not included in the initial list due to the AM+A's understanding that this project was seeking out the most diverse (orthogonal) range of collaboration practices within the client's international sites. We believe the spectacular increase in Chinese university training based on Western models has diluted the traditional Confucian orientation of young employees. By contrast, Japanese (and Korean and Singaporean) employees are expected to be more conservative in their communication style. In all cases, AM+A recommended specific locations with at least 200 employees to be researched so that the data collected would have strong statistical validity.

2.1 Data Visualization of the Recommendations

Using data supplied by the client and data that AM+A located, AM+A prepared a summary "visual table" comparing the international cities that AM+A recommended in comparison to the headquarters city. The values show the headquarters city presented with a value set at 1.0, and other values based on the data sources cited immediately below. The sequence emphasizes the cities that the client subsequently selected: Amsterdam, Bangalore, Singapore, and Tokyo, with data for Dubai and Shanghai shown additionally. The black rectangles show the maximum values for the criteria, and the white rectangles show the minimum values for the criteria.

The cities recommended provide strongly different culture attributes, which should make for effective, valuable studies of client corporate culture and differences and/or similarities in teamwork, which is the objective of the client's research. The sources of the data follow the table:

Cross-Cultural Comparisons with CA City

	CA City USA	Amsterdam Netherlands	Bangalore India	Singapore	Toyko Japan	Dubai UAE	Shanghai China
Power Distance, Hierarchy	1.0	0.95	1.93	1.85	1.35	2.00	1.22
Individualism	1.0	0.88	0.53	0.22	0.51	0.42	0.94
Digital Natives, Under 30 Years Old	1.0	1.07	3.60	0.40	1.13	0.73	2.00
High Growth, Newer Cities	1.0	0.43	2.18	1.03	0.17	2.70	1.94
Female/Male Gender Ratio	1.0	0.90	0.58	1.10	0.61	0.42	0.81

Aaron Marcus and Associates, Inc.

- *Power Distance:* With the exception of Shanghai, this data is drawn from the raw scores comparing 53 countries and national regions by Hofstede (1997). Due to his dependence on IBM data, Hofstede did not include China in his original study. As a result, the score for China is drawn from Schwartz's comparison of 39 cultural groups (1994) for Hierarchy, normalized from a 9-point scale in ratio to the US.
- *Individualism:* As with Power Distance, all scores for individualism are taken from Hofstede (1997) with the exception of Shanghai. That score is taken from Schwartz's score for Conservatism, reversed, normalized, and cross-validated against his scores for Egalitarian Commitment, Affective Autonomy, and Intellectual Autonomy.
- *Digital Nativit*y: This variable, defined as the client employees 30 and younger, is taken from employee demographic data.
- *High Growth/ New Cities*: This data is taken Urban agglomerations: Average annual rate of change (%), *World urbanization prospects: The 2007 revision population database,* available online from the United Nations Population Division (esa.un.org/unup).
- *Female/Male Gender Ratio:* Variable taken from client employee demographic data.

3 Analytical Criteria

3.1 Traditional Measures of Culture

Since the early 1980s, a large number of variables have been identified from factor analysis to explain differences between national and regional cultures. Some of these traditional measures include the following:

Hofstede's (1997) work dimensions (inc. Hofstede and Bond, 1988):	• Power distance • Individualism *vs.* collectivism • Long-term time-orientation (~ Confucian values)
Trompenaars and Hampden-Turner's (1998) cultural dimensions:	• Universalism *vs.* particularism • Communitarianism *vs.* individualism • Neutral *vs.* emotional expression (~ use of reason or feelings in relationships) • Diffuse *vs.* specific range of involvement • Achievement *vs.* ascription
Triandis' (2000) cultural syndromes:	• Individualism
Markus and Kitayama's (1991) notion of self-concept:	• Independent self-concept • Interdependent self-concept
Schwartz's (1999) cultural dimensions:	• Conservatism • Affective autonomy • Intellectual autonomy • Hierarchy • Egalitarian Commitment • Mastery • Harmony

Despite the large number of concepts, many overlap. (Different researchers factor analyzed different data sets, which led to divergent results.) In addition, the concepts defined by Hofstede, Trompenaars and Hampden-Turner, and Triandis are ranged along a single continuum from high to low or most to least (*e.g.*, power distance), while those measured by Markus and Kitayama and by Schwartz allow for greater diversity of behavior. According to Markus and Kitayama, individuals can hold independent and interdependent types of self-concept at the same time, but Hofstede theorized that people were *either* individualists or collectivists. (Note for comparison that Marcus and Baumgartner analyzed nine culture models by 11 authors, which elicited 29 dimensions that were rated by approximately 60 experts world-wide. From this set five "top" or most widely recommended dimensions emerged: context, technological development, uncertainty avoidance, and time perception. Data by country do not exist for all of these dimensions; they were not used for this paper.)

These traditional culture concepts are often associated with specific types of teamwork. In particular, Hofstede identified three major clusters of countries on the basis of ratings for power distance and individualism:

- High power distance/ collectivist: Mexico, China [Hong Kong], Singapore
- High power distance/ individualist: France
- Low power distance/ individualist: US, United Kingdom, Netherlands

People in collectivist countries with high power-distance (respect for authority) tend to be very dependent on their in-groups and rely on strong leaders who exercise moral authority. In addition, they may avoid direct confrontation (although they can engineer situations that let them apply rules to justify refusal). It is important to note that there are many historical variations of these values. India, with its tradition of close cooperation and mentorship between experienced and novice group members, is quite different from Tokyo, where managers are more separate from subordinates.

People in individualistic countries with high power-distance often develop bureaucratic systems that allow them to reconcile personal independence with absolute authority and centralized power. By contrast to collectivist countries, they may be more confrontational. People in individualistic countries with low power distance tend to work in loose groups and to treat others as peers; their allegiance to their in-groups is relatively weak so they find it easier to work with new people and outsiders. However, even within this cluster, there are historical patterns. Scandinavia is considered more egalitarian and "feminine"; work roles show less gender bias and groups look to achieve consensus.

There is much "surface validity" in Hofstede's categories, and he has been used in many studies of international business. However, when using Hofstede as a guide to study collaboration, it is important to remember that his data were collected from IBM in the late 1970s and early 1980s. Much has changed in global business since then, and much has changed in terms of the technologies and corporate strategies mandating teamwork. Hofstede was challenged by Michael Bond to develop a better measure for Asian countries. The result was a study based on Chinese cultural values, which looked at long-term vs. short-term orientation. As might be expected, China was ranked first of 23 countries, Japan fourth, and the US and the UK seventeenth and eighteenth. Several other client country sites ranked in between, with Brazil sixth, India seventh, Singapore ninth, and the Netherlands tenth. Trompenaars and Hampden-Turner's data are more recent (mid-1980s to late 1990s). Many of their culture concepts overlap with Hofstede's; others deal with notions of social systems first defined by Talcott Parsons in 1951 that are applicable to teamwork:

- Universalists use rules that apply to everyone and rely on procedural equity; particularists see social situations as more complex and tend to apply different rules to different types of people.
- Communitarians – like collectivists – place the needs and objectives of their group memberships ahead of their personal needs; individualists place their own needs first.
- People who display neutral affect tend to use reason as a basis for developing relationships; people who display a more emotional affect are more open about their feelings and expect emotional responses.
- People who have a specific range of involvement tend to view different business and personal contexts as separate from one another. They may recognize the authority of a manager at work but treat that manager as an equal when encountered at the mall. People who have a diffuse range of involvement with their work do not separate these contexts – a manager whose authority they recognize at work will retain that authority when met in a different situation.

Achievement and ascription refer to two methods used to assign group status and leadership. Status based on achievement is earned through an individual's personal accomplishments, especially work-related success. Status based on ascription is due to an individual's titles, gender, age, class, or education. Trompenaars and Hampden-Turner analyzed a somewhat larger group of countries than Hofstede but did not always measure the same group of countries for each culture concept. Using Trompenaars and Hampden-Turner's categories, the client's work sites were classified. Using these culture categories along with Hofstede's work dimensions began to show new differences between apparently similar countries like the Netherlands and UK, and Japan and China

A third culture theorist, Harry Triandis, believes individualism is the most important of all cultural dimensions for explaining behavior. However, he notes that this factor can be modified by a wide range of cultural "syndromes," like social complexity, tightness (ethnic homogeneity vs. heterogeneity), and hierarchy. Cultures that are less complex or highly homogeneous tend to be collectivist; most people will display more conformity and suppress individual expression. However, from our work with Japanese students and employees, it is important to remember that people in tight or collectivist societies continue to think of themselves as highly individualist; they just choose to put group interests first

Markus and Kitayama also support this notion of duality. They found that the cultural dimension labeled individualism vs. collectivism was too simplistic and unable to explain real-world behavior. As a result, they redefined the phenomena as two mutually-coexisting concepts: Independent Self-concept and Interdependent Self-concept. Their new variables better explain why American individualists give so broadly to charity while Malaysian collectivists do not. Strong membership in in-groups and an interdependent self-concept may prevent support for others in so-called out-groups. By contrast, people with an independent self-concept have a weak alliance to their in-groups and can see people in out-groups as being individuals just like them. These two types of self-concept may be important when looking at matrix management systems and the use of temporary work groups: in some cultures, longer-term groups may be more effective.

Finally, a last set of culture variables dealing with power and individualism was developed by Shalom Schwartz. His research was conducted in the late 1980s to mid-1990s; his values can stand alone but are also correlated. Conservativism is generally opposed to Intellectual and Affective Autonomy, Hierarchy to Egalitarian Commitment, and Mastery to Harmony. Conservatism (plus Hierarchy, and Mastery) roughly equates to Collectivism while Intellectual and Affective Autonomy and Egalitarian Commitment equate to Individualism. Schwartz's use of the term harmony refers to respect for the natural world, not group harmony. Thus, his values tend to highlight differences between countries like the US and China that focus on mastery for economic growth and areas like northern Europe seeking stable growth to preserve the environment. Schwartz also collected data from many countries and has tables listing values for his variables. Although he does not list rankings for the UK, UAE, Russian Federation, Saudi Arabia, or India, he does rank the rest of the client sites. Schwartz (1999) made some clear predictions about relationships between his cultural values and specific work behaviors in the following table:

Hypothesized Compatibility and Conflict of Culture Value Emphases with Dimensions of Work		
	Culture Value Emphases	
Dimensions of Work	**Compatible**	**Conflicting**
Work Centrality: Contrasted with leisure, community, family, religion	Mastery Hierarchy	Affective Autonomy Egalitarianism Harmony Conservatism
Societal Norms re Working: Entitlement *vs.* Obligation	Egalitarianism Intellectual Autonomy	Conservatism Hierarchy
Work Values: Power	Hierarchy Mastery	Harmony Egalitarianism
Work Values: Intrinsic (personal growth and creativity)	Intellectual Autonomy Affective Autonomy	Conservatism
Work Values: Extrinsic (monetary rewards)	Conservatism Hierarchy	Intellectual Autonomy
Work Values: Social	Egalitarianism Harmony	Hierarchy Mastery

This table suggests that Asian countries and northern European countries should be very different in terms of work centrality. In addition, like Hofstede, Schwartz recognizes France as having a unique blend of conservatism and intellectual and affective autonomy (roughly equivalent to high power-distance and strong individualism). Finally, he includes an element of environmental consciousness (relatively low for all countries with the client's sites) with his emphasis on harmony.

Summary of traditional measures of culture criteria: This review of culture theory highlights both the advantages and disadvantages of selecting sites for research on collaboration by using culture values alone. These values have strong impact. They do make it possible to hypothesize relationships to specific work behaviors and have been used widely in research. However, there is no one consistent set of variables, and many of the concepts were developed with data collected fifteen to thirty years ago. As a result, AM+A suggested using these culture values i*n conjunction with other criteria*, which seek to capture some of the changes taking place in corporate collaboration and technology-mediated teamwork.

3.2 Criteria: Technological Innovation and Globalization

The second set of criteria suggested for the client's study accommodates recent changes in the use of computer and communications technology and in global economics. There is a growing literature on the importance of recognizing generational differences, specifically, the rise of a generation of "digital natives," and the development of high growth cities and new forms of urban life.

As early as 1997, Johnston and Johal used culture theory to define the Internet as a new "virtual cultural region." Since then, the focus has shifted to the idea of *digital nativity*. Digital natives, or "millennials," have grown up surrounded by computers, mobile devices, video games, and the Internet. Older people have always used some other form of technology first; they are *Digital Immigrants*, held back by previous impressions of the "right" way to do things. Many claims have been made about the differences between digital natives and digital immigrants, such as the following quote by Prensky (2001) discussing changes in education: *"Digital Natives are used to receiving information really fast. They like to parallel process and multi-task. They prefer their graphics before their text rather than the opposite. They prefer random access (like hypertext). They function best when networked. They thrive on instant gratification and frequent rewards. They prefer games to "serious" work….But Digital Immigrants typically have very little appreciation for these new skills that the Natives have acquired and perfected through years of interaction and practice."*

To understand the full potential of new modes of communication, it is important to ask digital natives how they simultaneously work together and apart through social media. For the client to identify both current and future types of collaboration, sites with large numbers of new hires and digital natives should be visited. In the US, there has been some contention between digital natives and the baby boom generation; the baby boomers developed many of our traditional computer technologies but the new focus on social media often escapes them. They remain wedded to more structured media interactions, while digital natives thrive on instant availability and constant interaction. Outside the US, young Indian (and Chinese) college graduates have typically been the first to achieve computer literacy in their families. They join with fewer preconceptions about media and tend to see it in highly creative ways.

The rise of new types of cities is the second postmodern influence that the client should seek to capture in its study of collaboration. Throughout Asia, and parts of the Middle East, Latin America, and Africa, new cities are developing on the basis of new economic principles. In 1980, the estimated population of the United Arab Emirates was little more than 1 million; in 2009, it reached 5 million and, by the end of 2010, the government expects an additional 1.9 million for a total of 7.5 million (UAE Interact, 2010). For every child born in the UAE in 2009, 22 migrants arrived, making the country the world leader in immigration. Thus, 73.9% of the working age population (15-64) consists of non-nationals with 2.74 men per woman. The median age of the UAE is 30.1 years: 32 for men and 34.7 for women (CIA, 2010). Singapore similarly has doubled its population from about 2.4 million in 1980 to almost 5 million in 2009, but its age and gender demographics are now less skewed (Statistics Singapore, 2010). Current population growth is about 1% per year, mainly through immigration (15th in the world). The gender ratio is roughly equal, with a median age of 39 for both men and women (CIA, 2010), Both cities developed on the basis of trade, finance, and information rather than manufacturing, agriculture, or government.

Summary of technological innovation and globalization criteria: The client's professionals and staff currently working in such new cities are more likely to be globally diverse, young, digital natives who can anticipate future collaboration. Dubai currently epitomizes both the positive and negative effects of high growth and rapid urbanization. Investment in infrastructure has created a vibrant economy in the middle of the Gulf. However, gender ratios are highly skewed in favor of young men, and large numbers of expatriates are disconnected from traditional local culture. By contrast, Singapore followed a much more planned path to growth. Research there is likely to review more traditional Chinese attitudes to authority and collaboration.

3.3 Criteria: Equal Employment and Gender Ratio

Part way through our analysis, the client made available demographic data that allowed us to analyze the number of digital natives at various company sites. This data also made it possible to investigate an additional criterion: gender ratio.

Hofstede notes that countries differ on a continuum of Masculinity and Femininity, which refer to traditional gender roles and to attitudes associated with each gender. Men in masculine societies seek out jobs that require mastery; women take jobs that require nurturing others. Furthermore, men are expected to be assertive; women, modest. By contrast, men and women in feminine countries are less subject to rigid gender expectations. Men may be elementary school teachers; women, computer programmers. Both genders tend to be modest, avoid direct conflict, and respect (and expect respect) from others. Two aspects of a balanced gender ratio are important for collaboration. First, feminine societies tend to focus more on participation, persuasion, and consensus than masculine societies. Unions are included in company management and techniques like participatory design are popular. Masculine societies focus more on challenge, reward, and individual recognition. Master programmers are celebrated and proffered as role models for new employees.

Summary of equal opportunity and gender ratio criteria: The US has been fighting for equal employment in technical positions for decades. For a while in the 1990s, the gender ratio in computer science improved, but stereotypes of technology and of those who work in technological jobs ("geeks") have reduced the number of women attracted to such positions. As a result, AM+A suggests Amsterdam as a site for study of collaboration in feminine cultures and Bangalore and/or Dubai as sites for the study of teamwork in young, masculine cultures. According to the employee demographic provided by the client, the gender ratio in Singapore is slightly higher than in the company's headquarter site or in Amsterdam. Nonetheless, we believe that the number of women in technical positions may well be higher in the Netherlands and is more likely to provide an opportunity to analyze egalitarian styles of work.

3.4 Additional Considerations: Corporate Culture and the HQ Effect

During client meetings, challenges of collaboration between headquarters staff and the client employees in other countries were mentioned. This reference opened a discussion about corporate culture and its relation to national culture. The client employees are made aware of the corporate values. Many of these values reflect a classically US emphasis on personal fulfillment (empowerment, fun), mastery (innovation), efficiency (frugality), and continuous improvement. Teamwork is critical, but the *style* of teamwork tends to be more masculine than feminine. The consensus was that the client's corporate values provide a backdrop to discussions within the company about collaboration and that differences in power between headquarters and the regions are just as much, or more, of an influence.

4 Conclusion

This discussion presents AM+A's analysis of country cultures and specific cities, and their impact on values, attitudes, concepts, and behaviors related to teamwork. This analysis was used by the client to develop specific tools and techniques to improve collaboration, communication, and cooperation of multi-country teams.

References

1. Central Intelligence Agency. World Factbook (2010), http://www.cia.gov/library/publications/the-world-factbook/index.html (retrieved February 1, 2010)
2. Hofstede, G., Bond, M.: The Confucius connection: From cultural roots to economic growth. Organizational Dynamics 16, 417–433 (1988)
3. Hofstede, G.: Cultures and orgs.: Software of the mind. McGraw-Hill, N.Y. (1997)
4. Johnston, K., Johal, P.: The Internet as a ”virtual cultual region”: Are extant cultural classifications schemes appropriate? Internet Research: Electronic Networking Applications and Policy 9(3), 178–186 (1999)
5. Marcus, A., Baumgartner, V.-J.: A practical set of culture dimensions for global user-interface development. In: Masoodian, M., Jones, S., Rogers, B. (eds.) Proc. of CHI: 6th Asia Pacific Conf. APCHI 2004, pp. 252–261. Springer-Verlag, Heidelberg (2004), http://springerlink.com
6. Markus, H.R., Kitayama, S.: Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review 98, 224–253 (1991)
7. Prensky, M.: Digital natives, digital immigrants. On the Horizon 9(5), 1–6 (2001a), http://www.marcprensky.com/writing (retrieved February 1, 2010)
8. Prensky, M.: Digital natives, digital immigrants; Part II: Do they really think differently? On the Horizon, February 1, 2010. Retrieved from, vol. 9(6), pp. 1–9 (2001b), http://www.marcprensky.com/writing
9. Schwartz, S.H.: Beyond individualism/collectivism: New cultural dimensions of values. In: Kim, U., Triandis, H.C., Kagitcibasi, C., Choi, S.-C., Yoon, G. (eds.) Individualism and collectivism: Theory, method, and applications, pp. 85–119. Sage, Thousand Oaks (1994)
10. Schwartz, S.H.: A theory of cult. values and some implications for work. App. Psych: An Internat Review 48(1), 23–47 (1999) (retrieved February 1, 2010); EBSCO database
11. Triandis, H.C.: Collectivism vs. individualism: A reconceptualization of a basic concept in cross-cultural psychology. In: Verma, G.K., Bagley, C. (eds.) Cross-cult. studies of personality, attitudes, and cog., pp. 60–95. Macmillan, London (1988)
12. Triandis, H.C.: Culture and conflict. The Internat. Jour. of Psych., 35(2), 1435–1452 (2000); Reprinted Samovar, L.A., Porter, R.E., McDaniel, E.R. (eds.): Intercult. communication: A reader. Wadsworth Cengage, Boston (reprinted 2009)
13. Trompenaars, F., Hampden-Turner, C.: Riding the waves of culture: Understanding cultural diversity in global business, 2nd edn. McGraw-Hill, New York (1998)
14. Statistics Singapore (2010), http://www.singstat.gov.sg/stats/themes/people/hist/popn.html (retrieved February 1, 2010)
15. UAE Interact UAE population to rise in 2010 to 7.5million (January 25, 2010), http://uaeinteract.com/news (retrieved February 1, 2010)

Scalability of UX Activities in Large Enterprises: An Experience Report from SAP AG

Boris Hecker and Matthias Berger

SAP AG, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany
{Boris.Hecker,Matthias.Berger}@sap.com

Abstract. If a large company like SAP is selling business software, they face several challenges when you try to introduce new processes or change existing ones to increase the usability of your software. You need to get the right people with the right skills, you need to have the budget to run dedicated activities like user research or usability tests, and you need support from other internal organizations and management. The organizational part is relatively easy to address; a much harder challenge is to change people's mindset and thinking. It is not enough to introduce a new process and tell everyone to execute on that new process. You need a large commitment from the people executing the processes in order to change their behavior and the way they do things if such a large change is to happen. At SAP we are in the middle of such a change process and want to show you which challenges we have encountered and how we have dealt with these situations. We will also deep-dive into a concept which we call "User Experience Advocates (UxA)".

Keywords: Usability, Scalability, User Centered Design, Business Software.

1 Introduction

To develop high-quality business software, you first need a large variety of skills and people to make the product a success. It is necessary to have a development architect to take care of the technical parts, a business expert to take care of the business processes and customer requirements, and you need a user experience professional to cover the end-user requirements and the UI design.

In addition to having the right skills and the right people in the project, you also need an environment where these experts can perform at their best.

What does that mean?

This means that management needs to set the stage for the project in terms of time, budget, and resources. If you have enough resources and money but not enough time you won't be successful with the product. If you have resources and time but no money it is the same; you will fail. Also, if you have time and money but no resources, you will fail. If you want to develop a successful product, you need to have these three things in place.

In a large enterprise that can easily run 1000 projects a year, you need the resources and money to be able to scale with that number of projects. You don't need

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 425–431, 2011.
© Springer-Verlag Berlin Heidelberg 2011

to care so much about the time in this case, as you will run most the projects in parallel. But the money, and more important, the resources are the most critical part.

SAP has a User Experience (UX) department which is responsible for the usability of the whole SAP product portfolio. The User Experience department at SAP was founded in 2005 out of disparate groups and decentralized areas. UX has about 200 members from which we have approximately 38 people working as so-called user interface designers in the projects of the On-Premise Business Suite. We will focus on that area as here we started the pilot to introduce the concept of the User Experience Advocates.

Thirty eight people are not nearly enough to support all necessary projects and to help them to develop an easy-to-use application. The big question is now how the company can scale the expertise and the knowledge available in the UX area with the few experts available!

This paper will give you a quick introduction to SAP and the UX department. After that we will discuss the general understanding of usability in the company and how the overall UI literacy has increased over the years. After that you will learn about the four most common areas of improvement within SAP and finally, you will learn how these areas could be addressed by implementing the concept of the UX advocates.

2 SAP

As the world's leading provider of business software, SAP (which stands for "Systems, Applications, and Products in Data Processing") delivers products and services that help accelerate business innovation for our customers. We believe that doing so will unleash growth and create significant new value – for our customers, SAP, and ultimately, entire industries and the economy at large. Today, customers in more than 120 countries run SAP applications – from distinct solutions addressing the needs of small businesses and midsize companies to suite offerings for global organizations.

Founded in 1972, SAP has a rich history of innovation and growth as a true industry leader. SAP currently has sales and development locations in more than 50 countries worldwide and is listed on several exchanges, including the Frankfurt Stock Exchange and NYSE under the symbol "SAP."

2.1 SAP and User Experience

In the past SAP only had a small group with a few experts in the area of HCI. This group was on a regular basis centralized and decentralized due to reorganizations. Their official job role was developer or business expert but there was no official role like user interface designer. Due to that the work was also not valued as these people maybe deserved it as they were the very first ones who fought for the usability of the applications within the company.

In 2005 Dan Rosenberg joined the company and started to pull all HCI experts within the company together to form a central User Experience organization. UX was brought to life in 2005. The mission statement of SAP UX is:

"The SAP User Experience group's mission is to improve the user experience of SAP applications"

Today, after UX was founded, the department has about 200 experts in different areas with an education background in e.g. visual design, psychology, communication design, media-system design, computer science and more. The experts are spread across 14 locations worldwide. SAP UX is coordinating the activities with the user-centered design process.

3 Challenges

With the long history of SAP as a fast growing company in the 70s and 80s, we face several challenges which have their root cause in the past of the company.

1. Development-driven decision-making. The company was, from a historical perspective, driven by features and functions. Features and functions were the selling argument back in the 70s, 80s, and 90s, but this has changed! The business departments of our customers have more and more power and influence purchasing decisions. The end-user, the usability, and, in the end, the whole user experience of a product gets more important than the features and functions. The scale moves from a technical-driven way of building software applications to end-user-driven way.
2. Limited UX Resources. The number of available UX experts is not enough that we can staff every project which has a need for such skills. Often projects do not get the support they really need.We also face the challenge that one UX expert is supporting more projects than in parallel. As a consequence there is the need to split up the time to support multiple projects. This results in a decrease in the usability of the products. The ratio between developers and user interface designers at SAP was in the past approx. 150:1! Today it is about 100:1. A good ratio would be 10:1.
3. User interface designers were more fire fighters than designers. As the general UI literacy was not that high, most people had no understanding of what impact it has to just develop something without knowledge of the end user and how this person is doing a task. In the past it worked very well without that knowledge, but today it is absolute necessary to consider these points. A common situation is that at the end of a development project the project lead steps into the office and asks for a UID to make the UI usable. They are often not aware that there is a huge difference between usability and visual design. This demonstrates the lack of awareness that leads them to fail. The projects often just do not plan any UX activities upfront.
4. UI Design Literacy level of the company was low. The understanding of the value that UX can bring into the product was not there in the beginning. There was no training available on that topic. It was not integrated in the overall development lifecycle of SAP products. No common UI guidelines or style guides were available. Five years ago SAP had over 21 different UI technologies used across their products, all with different capabilities, interaction behavior, and visual design. Every development team was able to introduce new UI technologies. There was no governance in place to control this.

Three points were identified, that needed to be addressed.

1. Continue to raise the SAP corporate UI design literacy level and attention to design detail
2. Scale direct UI knowledge, responsibility & ownership inside product development teams
3. Scale professional UX design and user research skills for critical areas.

4 UI Literacy Level

To further increase the awareness of the importance of usability and end-user-centered interfaces in business software, UX needs to continue to educate the whole company in this area. UX also needs to show the value of building highly usable software to the internal stakeholders.

It is actually very challenging trying to measure usability. Upper management is often interested in proven figures they can compare. Usability also depends on the person using the software, their experience and the mental model of this person. Hence, it is not an easy task to get such figures.

It is important to understand that the management should be the first area to address these topics. It is necessary to convince them to help to increase the UI literacy level in the company. Without management support it is nearly impossible to be successful.

It is often the case that during a training, people come to understand all aspects of the material, but when it comes to the point to apply the knowledge in a daily life situation, many people fall back into old behaviors and are not able to apply the learned things. The purpose is necessary to break this cycle and change the behavior and attitude, and this has to be done by everyone themselves. It will not work if the management says from now on you do things differently and each and everyone still does the task as they did it in the past.

In our case, to drive this change, we constantly raised the UI literacy level in the last 6 years more and more.

The foundation of UX and later the introduction of the UX advocate concept were two milestones with a major impact on how the company gained momentum to trigger further activities. Strong improvements contributing to this change were…

1. Establishing the UI governance model to control the usage of UI technologies
2. Set up a UI Guideline team to centrally push the UI guidelines into the projects and ensure consistency across the products
3. Incorporating the user-centered design process into the product development process
4. Improving the UX infrastructure by offering training and workshops
5. Setting up a dedicated role in HR to support and reflect the colleagues working in the UX area.
6. Running and promoting successful projects – internal and external.

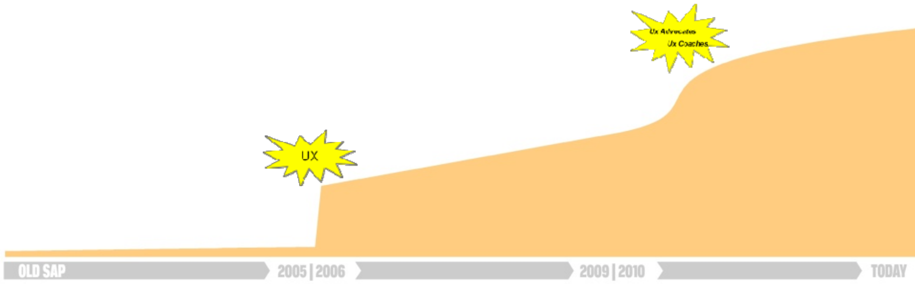


Fig. 1. The increase of the UI literacy level over the last 6 years with the 2 major events that had the biggest impact

In the long-term, it is very important to raise the UI literacy level within a company but it will not help you to scale your UX activities into the organization in the short term. In the short-term, you still have the problem that you do not have enough resources to support all upcoming projects in a sufficient way. The majority of the projects still have not included UX activities in their overall project plan.

To change the way people do things it is absolutely necessary to make the value of doing things in a different way transparent and easy to understand. The best way is to put the people in the position to do UX tasks by themselves.

In that situation the idea of the user experience advocates was born.

5 UX Advocates and UX Coaches

The idea is that every team needs to take on responsibility for their developed UIs, make sure they fit to the needs of their end users, comply with the UI guidelines, and follow the basic design principles. The projects cannot rely on the central UX team to "make the UI usable" and fix the UI at the very end of the development cycle because nobody took care of it in an early stage. The user experience professionals are not there to act permanently as fire fighters for usability issues. Some members of the development teams themselves must supplement their skills by also taking on the role of an Ux Advocate, supported by the central UX infrastructure, acting as a bridgehead into it.

*How to bring the UX responsibility into the development teams?*A company could just give an internal command to do it. But would this work? Of course not, because you need the mindset, the skills and the commitment of the individual people to do this. By just commanding this, people will not change how they do the work.

The first step is to convince the management to free up some of their resources and to transform them into advocates for the user experience of the products. This means that if a developer takes on the role of an UxA, the project will lose development capacity. Here, management needs to support this by freeing up time for people who want to take on this role and not force their employees do those tasks on top of their development tasks. It is also very important that the management respect and value the work of an UxA in their area of responsibility.

Fig. 2. The three dimensions of UX knowledge. The UX core, with the experts who coach and support, the UX advocates in the different projects, and finally all other people involved.

Table 1. Description of the tasks and responsibilities of an UxA and the support an UxA will get from the central UX organization and the UX coaches.

Tasks	Responsibilities	Support
▪ Plan, track, and report on UI-related tasks for sprint backlogs ▪ Assist with user research (e.g. Create User Profile, requirements elicitation, formative usability tests) ▪ Distill business and user requirements into use cases and user stories ▪ UI design including interaction design, navigation, and flow ▪ Create prototypes (sketches & wire frames) ▪ Identify new interaction design patterns (if applicable) ▪ Conduct design review sessions with project stakeholders and customer representatives ▪ Conduct UI quality reviews to ensure that implemented UIs match UI mock-ups and conform to UI guidelines ▪ Support in UI-related customer engagement activities (e.g., formative usability tests, customer UI walk-throughs).	▪ Create User Profiles ▪ Create Use Cases Create UI design prototypes ▪ Ensure UI quality and guideline compliance ▪ Consult with Ux Coaches	Coaching on ▪ Planning of UCD activities ▪ Selecting the appropriate user research methods ▪ Creation of use cases ▪ How to get from use case to design ▪ Prototyping (e.g. paper-/Visio mockups) ▪ Creation of guideline compliant UIs ▪ Validation of UCD deliverables (e.g. use case, mockups) ▪ Usability testing ▪ Application of Product Standard Usability

The second step is to talk to the people and identify those who wants to do the job and take on the challenges and opportunities of such a role.

The third step is to educate and motivate these people so that they are able to deliver high quality and end-user oriented UI designs.

It is important to have a core of UX professionals growing a pool of people with the right educational background, skills and experience in the HCI/UCD area.

Around that core, the education of the UxAs takes place to act as an advocate for UX topics within their area and to collaborate with the UX professionals who act as a coach for the UxA.

The remaining stakeholders composes of all employees not in a role of an UX professional or an UxA who need a basic understanding of the HCI/UCD area in order to accept and facilitate certain processes and methods to reach a usable product. A prime example is the project lead who needs to understand when to look out for a designer or a person who can do the design task in his project.

6 Conclusions

To scale UX activities in a large enterprise company you need to take on the challenge to:

- have a good awareness in the area of UCD/HCI in the whole company to avoid time consuming discussions about the value of user research, conceptual design and the need to follow the basic design principles. And most important to be able to run UX activities in the projects in a pragmatic and simple way.
- make the project self responsible for their UI by nominating someone in the team to actually do the user interface design and take care of it. This person will get strong support from the UX professionals and will work in a very close collaboration with the UX coaches. This person also needs recognition from management.
- provide an infrastructure to educate the company and especially the UxAs to do the job by offering training, workshops, and infosessions and by having a very close coaching model in place.

The biggest challenge is to make the change happen. To make people do things differently and change their working behavior. Everybody should understand that such a change does not happen over night; This is really a long-term plan and you need a lot of patience and endurance to reach the goal. The process will pay out in the end and is absolutely worth going that way.

With a variety of activities and the concept of the UX advocates we area able to address all these issues, but it is certainly, still a long journey to reach a UI literacy level in the company that it is simply matter of course to run UCD activities in the projects.

Training Designers of Real-World Products: Scenario Approach in Industrial Design Curriculum

Yu-Chun Hsu

Instructor, Dept. of Industrial Design, Chang Gung University, Taiwan
hidrtommy@gmail.com

Abstract. This paper aims to explore the use of scenario approach in training curriculum for industrial design students. Over-reliance on internet and ignorance of user's need are critical issues to be addressed among design students. It is of interest to incorporate Scenario Approach to education content to overcome students' failure to grasp the key points of a design. We investigated differences in views on scuba lighting equipment between design students and expert users. Two groups of 23 participants respectively from the Professional Association of Diving Instructors (PADI) and an industrial design department were recruited. Questionnaire was used for collection of data in this study. Narrative statistical analysis and independent variable t-tests were used to interpret the raw data. The findings indicate that design students and expert divers place their primary focuses on completely different factors. Overall, the students put a great deal more emphasis on appearance of lighting equipment than did the expert divers. Divers rated multiple functionality and weight among the most important factors in design. The students sampled largely overlooked these two key points. Results of this study are critical to industrial design instructors. It is suggested that students must come to understand usage patterns and cognitions of users prior to preparing a design project. Students can gain such understanding through application of Scenario Approach combined with real-life experience of the equipment. In addition, due to the wide range of materials available on the Internet, and their unregulated nature, the quality of information that design students access online varies wildly. Furthermore, because the Internet is convenient, online information may be overused and misapplied by design students.

Keywords: scuba lighting equipment, PADI, Scenario Approach, Internet.

1 Preface

Design students receive massive information on internet as assistance or stimuli for their design. It is recognised that students simultaneously over-rely on and manipulate other's ideas from the Internet. It is of interest to transform this phenomenon in the final assignment required by the curriculum of design methodology, and then, more importantly, reshape their existing positioning in design.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 432–441, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Initially, scuba lighting equipment was chosen as a topic for this curriculum, which was appropriate for differentiating user's experience, and facilitating the subsequent evaluation. Participating students were introduced with scenario approach that effectively help ideation of students and allow student to visualize the potential problems that a user would encounter. Moreover, such approaches avoid improper information assumption and product cognition.

The aim of this paper is to help students recognize that design is an interactive creation process between perception and experience, and that the validity of an idea may reveal designer's information approach in his/her long-term memory. And thus, once one's design experience meets a certain degree, he/she will be able to decipher the information of user's experience, and in the meantime, transform the elements in design creation (Hsu1996). Questionnaires were designed after reviewing relevant literatures and internet tests were then prepared. Students were guided for the steps of scenario approach and the assessment list.. The researcher formally conducts the questionnaire of the Internet, after the pretest is re-modified by Mr. Hwang. The data—after collecting, analyzing and recording—will be compared and discussed. Finally, the paper attempts to proffer a research conclusion and suggestions for future researchers.

2 Literature Review

Scuba lighting equipment is main lighting equipment for the scuba divers along with one or two smaller lighting equipments, in terms of its lighting and capacity. The smaller one will replace the main one, when the main light equipment is out of power or broken. In addition, regarding conventional battery, user has to install a new battery before doing exercise whereas rechargeable battery is charged before diving (Lin, 1994).

Beyer and Holtzblatt (1998) divide user's experience into fives models: 1) flow models; 2) sequence models; 3) cultural models; 4) artifact models; and 5) physical modes. The five models will, through user's possible problems and messages encountered in the five modes, reveal user's experience model.

Flow model: we can discern the relationship involved in this activity, including the intertwining interaction among people, things, and objects.

Sequence models: the framework of the model aims to achieve its goal. From beginning to end, the sequence of steps—encompassing the contingent events, which are demonstrated as "trigger," "steps," and "order breakdowns"—aims to achieve one goal of things.

Cultural models: this reveals the impact of different organization or roles upon the whole culture. The individual organization or role may achieve the whole structure of a certain event or one small minor impact. From the cultural impact on the events, we can understand how it affects one's values of notion.

Artifact models: we can analyze the actual object and then detect the use frequency and location of the object from the framework of the models. Then, their relationship can be discerned via the succession of time.

Physical models: the arrangement of the objects reveals the interrelating relationship between object and humans. This relationship may help clarify the issues of environmental arrangement.

The scenario approach was early applied to human-computer interaction (HCI). One typical example of it is Richardson Smith, who was in the ID TWO design company in England and America, applied scenario approach to develop the panel design of the printer. At the very beginning, one has to observe the scenario of the product, and then guide him/her into the human-computer panel design. Then, this approach was more and more applied to the design of the product (Kelley, 2001). The sequence of scenario approach, as IDEO (1992) proposed, can be divided into five steps:

1. Understand
2. Observe
3. Visualize
4. Evaluate and refine
5. Implement

Scenarios approach suggests that one, in the development of the product, simulates a using situation through imagination, including user's characteristics, events, and the relationship between the product and the environment. This approach aims to analyze the relationship between humans and product via the using situation. By the visual and actual experience, this approach guides and participants in the product development. More importantly, it aims to invent any new product idea, judges whether the idea goes with the design theme, and evaluates whether the idea of the product reaches user's potential need. This is, so to speak, the user centric design (UCD) (Tu, 2005).

3 Methodology

3.1 Procedures

The researcher attempts to connect scenario approach to user's experience model, as Fig.1 demonstrates. Thus, we can apply this model to the relationship among people (students and experts), things (scuba diving), and objects (scuba lighting equipment). Moreover, the observational sphere—steps, procedures, stimulation, and decomposition—can be approached to strengthen the depth and width of observation. We can make the model of the product visually vivid by use frequency, use position and duration of time. After the three steps, the assessment and refinement will be made to the Internet questionnaire and expert's evaluation; the questions of the Internet questionnaire are, of course, what the researcher aims to inquire and investigate. Then, the differences between the production and cognition will be evaluated by the experts.

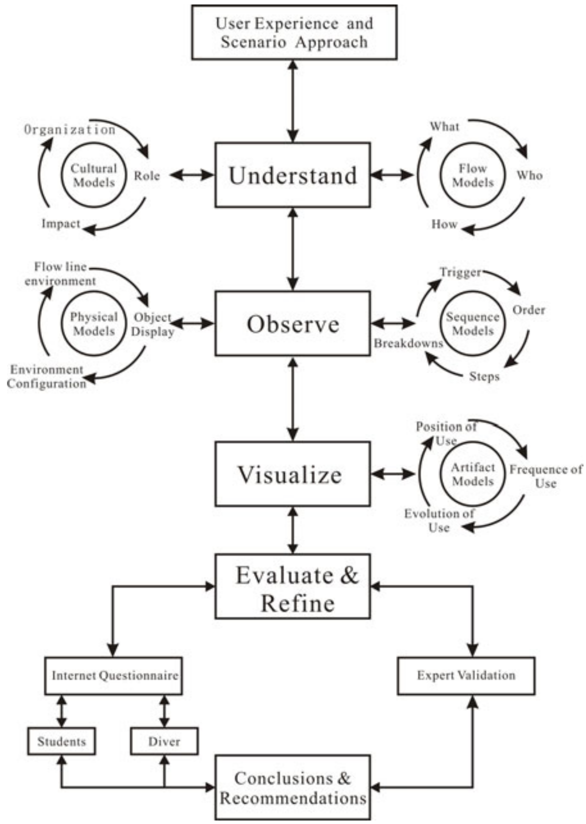


Fig. 1. Research Process

3.2 Design of the Questionnaire Content

After the topic of the research is ensured, the researcher begins to design the content of questionnaires, in which, as the research wishes, the interviewees may offer valid and beneficial information for the research. The content of the questionnaire includes the scuba diver's information, gender, age, address, email, education background, job, frequent regions of scuba diving, frequency of scuba diving, scuba diving system and the levels of scuba diving license. The above information will facilitate the researcher to carry on the further tracing and feedback. Meanwhile, the students of design fill in the above information. This will be used to do t-test analysis between the two groups for their differences.

3.3 Limit of Participants

There are two groups for the research: one is the group of 23 students of design, and the other is the group of scuba diver. The limited numbers of participants will be the characteristics for the research due to the fact that the number of professional scuba divers may reach the real group subject.

4 Design of Internet Questionnaire

4.1 Design of the Databank Program

The design of questionnaires has to go beyond the traditional questionnaire, because the group is very unique (Chu, 2002). The researcher is able to receive the information via Internet, instead of the written form of questionnaire. First, the questionnaire content is made with a refined and beautiful webpage. Through the inscription of MYSQL (Matthew and Stones, 2003), each item clicked by the testee will be transmitted to the server of researcher's computer, and then stored in the databank (Fig. 2).

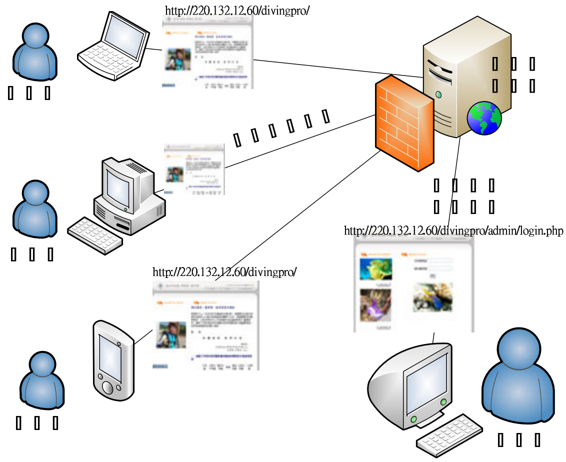


Fig. 2. Design internet questionnaire and indicate data links with databank

The strength of the Internet questionnaire is that the taste is able to fill in the questionnaires beyond the limit of time and space. This is very convenient for the statistical investigation since scuba divers' diving regions are sometimes different.

Then, after the information of the databank is established, the statistical data will be operated via the computer, instead of the traditional manual collection. This is not only time-saving, but also ecological friendly.

The webpage can be accessed at < http://220.132.12.60/divingpro/ >, which provides the expert instructors' discussion and scuba divers' revised questionnaires. In this webpage, the head instructor gave a prologue for this webpage in hope that this website provides a convenient and sustaining site for the discussion and pretest of the design of scuba diver's equipment (Fig. 3).

Fig. 3. Questionnaire content and webpage design

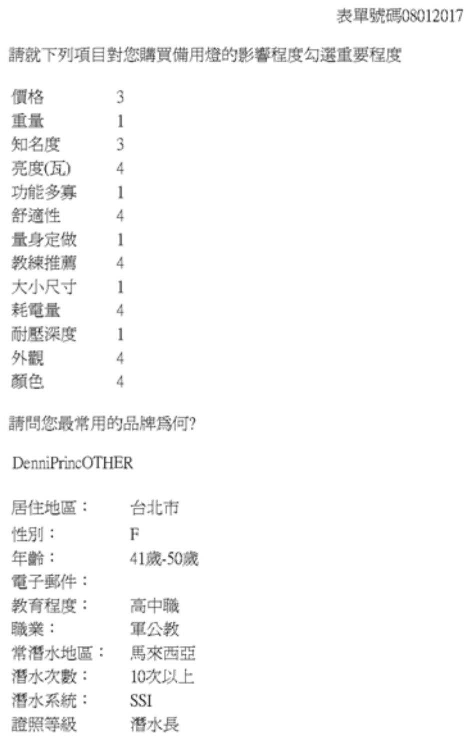
4.2 A Site of Data Collection

The items clicked or sought by the tastes on the webpage can be explicitly seen and meanwhile, the results of taste's feedback can also be demonstrated. In addition, the interaction between the MYSQL and the webpage can also reach the need of function.

Fig. 4. The permitted ID number and password

The website of administers’ log-in is at: <http://220.132.12.60/divingpro/admin/login.php>.

The tastes who have the permitted ID number and password (See Fig 4) are able to fill in the questionnaire and see the results of their questionnaire in the databank (See Fig. 5). The administrator can, through the convenient and advanced technique of Internet, be informed by any on-going information received.


表單號碼08012017

請就下列項目對您購買備用燈的影響程度勾選重要程度

價格	3
重量	1
知名度	3
亮度(瓦)	4
功能多寡	1
舒適性	4
量身定做	1
教練推薦	4
大小尺寸	1
耗電量	4
耐壓深度	1
外觀	4
顏色	4

請問您最常用的品牌爲何?

DenniPrincOTHER

居住地區：	台北市
性別：	F
年齡：	41歲-50歲
電子郵件：	
教育程度：	高中職
職業：	軍公教
常潛水地區：	馬來西亞
潛水次數：	10次以上
潛水系統：	SSI
證照等級	潛水長

Fig. 5. The questionnaire and see the results

5 Analysis

5.1 Result

The 23 (college) sophomore testees are given with the paper questionnaire, while the expert scuba divers are with Internet questionnaire. The Internet questionnaires received are 23. The result of each variable of the questionnaire is in detail demonstrated as follows:

Table 1. Sigma of each variation

	Student's group	Scuba diver's group
Price	5.13 (1.632)	5.52 (1.928)
Weight	5.70 (1.02)	4.61 (1.994)
Appearance	5.43 (1.037)	3.87 (1.890)
Multifunction	5.13 (1.014)	4.83 (4.370)
Comfort	5.91 (0.848)	5.70 (1.063)
Lighting	6.04 (0.825)	6.48 (0.593)

5.2 The Analysis of Independent Variable T-test

The ANOVA statistic analysis of SPSS is conducted in this research; the significant standard P value is 0.05; t-test is adopted to analyze the six variations (which are, price, weight, appearance, multifunction, comfort, and lighting) between the college students and expert scuba-divers while purchasing the lighting equipment of scuba diving. The results of the analysis are mirrored as in Table 2.

Table 2. Result of Independent sampling t-test

	Levene test		T-test
	F-test	significance	Significance (two-tailed)
1_price	0.312	0.579	0.461
2_weight	13.347	0.001	0.026*
3_appearance	3.272	0.007	0.001*
4_multifunction	1.535	0.222	0.396
5_comfort	3.510	0.068	0.447
6_lighting	0.089	0.766	0.046*

Note * suggests there is significant difference, P<0.05

From this, we are able to discover that the values of the variations such as price, appearance, multifunction, comfort, and lighting are more than 0.05, suggesting that the variance in the five variations is equal; the values of variation such as weight is 0.01, which is less than 0.05, suggesting that the variance of the variation is not equal.

From significance (two-tailed), we can discover that the P value of the three variations in price, multifunction, and comfort don't have significant difference with one another. However, for their P value is less than 0.05, the two variations of appearance and lighting have significant difference. Moreover, because its P value is less than 0.05 (which is 0.026), the variation of weight have significant difference with the invariant variance.

The analysis of t-test reveals that the three variations—which are weight, appearance, and lighting—have significant difference.

5.3 Result Analysis

The mean of weight in the college student's group is 5.7, while that in the scuba diver's group is 4.61. This difference may primarily derive from their difference of experience. Most of the college students did not have any experience of scuba diving, and they failed to know the fact that the scuba diving light doesn't almost have weight in water. In contrast, the expert scuba divers, with their lived experiences, may less likely take the weight into consideration. Moreover, because the light is not usually big, it doesn't cause any difficult for those who want to carry with it.

The mean of appearance of the student's group is 5.43, while the mean of the scuba diver's group is 3.87. The differences between the two groups may primarily account for the fact that the students, whose educational background is in design, may be much more susceptible to the appearance or the design of the object. They, of course, may greatly take the appearance into consideration. However, the professional scuba divers deeply know that it will be very dark in the depth of 20 meters of the underwater; one can hardly see in that depth. Thus, their primary concern for buying the scuba diving equipment is multifunction. It'll be very convenient and safe once the equipment is combined with multiple functions, especially if they encounter any urgent or dangerous situations. Moreover, they don't have to carry much equipment with themselves under the water. This difference may proffer an important indicator for the future designers.

The mean of lighting in student's group is 6.04; that in scuba diver's group is 6.48. The difference may, due to the medium of water, explain the fact that lighting will be more reduced in water rather than in land. Thus, generally speaking, the two groups will consider the lighting of the scuba diving equipment. But more specifically, the scuba diver's group will consider lighting more than the student's group. Moreover, the other reason to account for this difference may primarily rely on the fact that lighting is a keenly and desperately desired in the dark and deep underwater (Richardson, 1995). This can hardly be discerned or comprehended by the student's group. People are phototoxic oriented; they will feel the sense of secure whenever there is light. The place with light is, after all, more sought and desired than the place with darkness. That lighting has significant difference between the two groups can be another indicator or variation for the design of function.

6 Conclusion

We discover that the two groups have, as regards user's experience on the lighting aid of scuba diving, significant difference in terms of weight, appearance, and lighting.

Most design students have been trained with simulations and interviews of scenario; however, most of them, without real experience of diving and other group's situation and cognition, have missed the focal point of the product designed in the final stage.

Most students are accustomed to searching for information via Internet and use it without further consideration or evaluation. This phenomenon is explicitly reflected on the overlapping of student's design works. Worse, some works, though made by different groups, are completely the same. Then, at the beginning of the idea, we can clearly see students' misjudgment on the information, though some will be pinpointed by the experts. However, the differences of the experience will ensue and develop.

The result of the Internet questionnaires is simultaneously done with the data of the server. This is, beyond doubt, an interdisciplinary integration among design, computer science, mechanical engineering, and scuba diving; this offers a helpful and convenient site for the data collection, discussion, and cooperation for each field.

References

1. Beyer, H., Holtzblatt, K.: Contextual Design, Defining Customer-Centered Systems, pp. 89–123. Morgan Kaufmann, USA (1998)
2. Lin, K.C.: Recreational Scuba Diving. M. A. thesis, National Sport University (1994)
3. Hsu, C.L.: Research of Application to Conceptual Creation from User's Experience. M. A. thesis, National Chiao Tung University (1996)
4. IDEO. Innovation design and interaction design (1992); In: Kelley, T. (ed.) The Art of Innovation, A workshop for ACER 27, Doubleday, New York (2001)
5. Neil, M., Stones, R.: Instruction Manuel of MySQL, pp. 49–58 (2003)
6. Richardson, D.: PADI Open Water Course. In: Li, S.H., Hua, C.S. (Trans.) Wen Shih Commercial Press, Taipei (1995)
7. Tu, L.C.: Creative Learning for Designing Cultural Products by Scenario Approach: A case study on Taiwan aboriginal garment. National Digital Archives Program (NDAP), National Science Council (NSC), Taipei (2005)

Why Enterprises Can't Innovate: Helping Companies Learn Design Thinking

Jon Innes

UX Innovation LLC, 330 Mission Bay Blvd., Suite 506,
San Francisco, CA 94156, USA
jinnes@acm.org

Abstract. This paper describes the challenges of introducing design and innovation best practices into large companies focusing on enterprise software products. It proposes a theory for why existing companies tend to focus primarily on technical development factors while largely ignoring ease-of-use of the resulting solutions, and yet have been commercially successful to date. It also proposes that the market circumstances that have enabled this situation are likely to change as the market continues to mature. It then suggests methods for existing vendors to adapt to those changes and outlines the risks of not doing so.

Keywords: Design, innovation, user experience (UX), user interfaces, enterprise software.

1 Introduction

In the middle of the twentieth century, noted author Rudolph Flesch, wrote *Why Johnny Can't Read* [1] addressing the fundamental flaw in educational training that was preventing young children from mastering basic reading skills. As a result, significant changes were made in early educational programs around the globe helping many children gain basic literacy skills. Today, human computer interaction (HCI) professionals face a similar challenge in teaching corporations how to become more innovative via user-centered methods.

Roger Martin of the Rotman School of Management at the University of Toronto has written a book *The Design of Business: Why Design Thinking Is The Next Competitive Advantage* [2] that calls for all business leaders to reconsider the current organizational paradigm in favor of optimizing for what he calls "*Design Thinking*". The similarity to Flesch's work is striking. Martin is not alone in his views, Silicon Valley author Michael Malone [3], IDEO CEO Tim Brown [4], and other's have similar suggestions.

When surveyed, the consensus of these leading thinkers is that today's fast moving global business environment requires businesses to master two key types of innovation for their long-term survival:

1. Refining innovation, also known in HCI circles as N+1 innovation [5]
2. Disruptive innovation, aka "breakthrough innovation" [6]

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 442–448, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Without refining innovation, existing market leaders risk being left behind by competitors who can refine a product or service for a market faster than they can. Without disruptive innovation, companies serving markets are at risk of being displaced by new competitors who define entirely new products or services, making existing business models obsolete. Both types of innovation are at the essence of HCI work focused on shaping technologies to map to human needs.

2 Transformation Requirements for Innovation

As the above-mentioned authors and many others have noted, the process of helping companies improve their design thinking skills is not a simple one. It requires deep changes in the organization's culture and structure. Most large companies value process efficiency that drives out variation and ambiguity.

Variation and ambiguity are key elements in any creative endeavor. As such, creating an environment that encourages innovation requires changes that are counterintuitive to most managers who have developed their skills entirely while working within large enterprises. Leading such a change in mindset at most large companies is akin to attempting to change the DNA of a living organism. These are endeavors with high failure rates. Transformations of this sort remain more art than science.

Business books are full of stories of failed attempts of enterprises to develop "design-driven innovation" programs. What are the lessons we can learn from these past efforts? If we hope to codify the insights and gut feelings of early thought leaders into repeatable best practices, analysis and refection is required. Here are some of the factors to consider as HCI professionals helping organizations attempting to transition to a state of user-centered design and innovation competency.

2.1 The Resources, Processes and Values Framework

Clayton Christensen's popular series of books on innovation provide a very comprehensive analysis on what prevents companies from innovating. His thesis is well supported by numerous case studies.

Christensen's first book, *The Innovator's Dilemma* [7], introduces a framework that is very applicable to the enterprise software space. That framework, which emphasizes three key components *resources, processes, and values*, explains why many HCI efforts fail, especially within large established companies developing what is known as Enterprise Resource Planning (ERP) in the software industry. This is somewhat ironic given ERP products are intended to help companies manage processes and resources, and yet the ERP vendors struggle with implementing design processes.

2.2 The Values Problem

Most ERP companies do not value user-centered design or the benefits it provides, a good user experience (UX). Actually to clarify, most ERP companies fail to understand the value of UX in ERP products. They view good UX as unnecessary for the sale of ERP solutions. Why is this? Norman's [8] theory, as depicted in Figure 1 below, explains this if one concedes that the ERP market is still in an immature state relative to the market for consumer electronics or eCommerce sites. This is a reasonable position given that ERP vendors mainly compete against IT teams developing "*custom software*" that fail to deliver on over 90% of their projects [9].

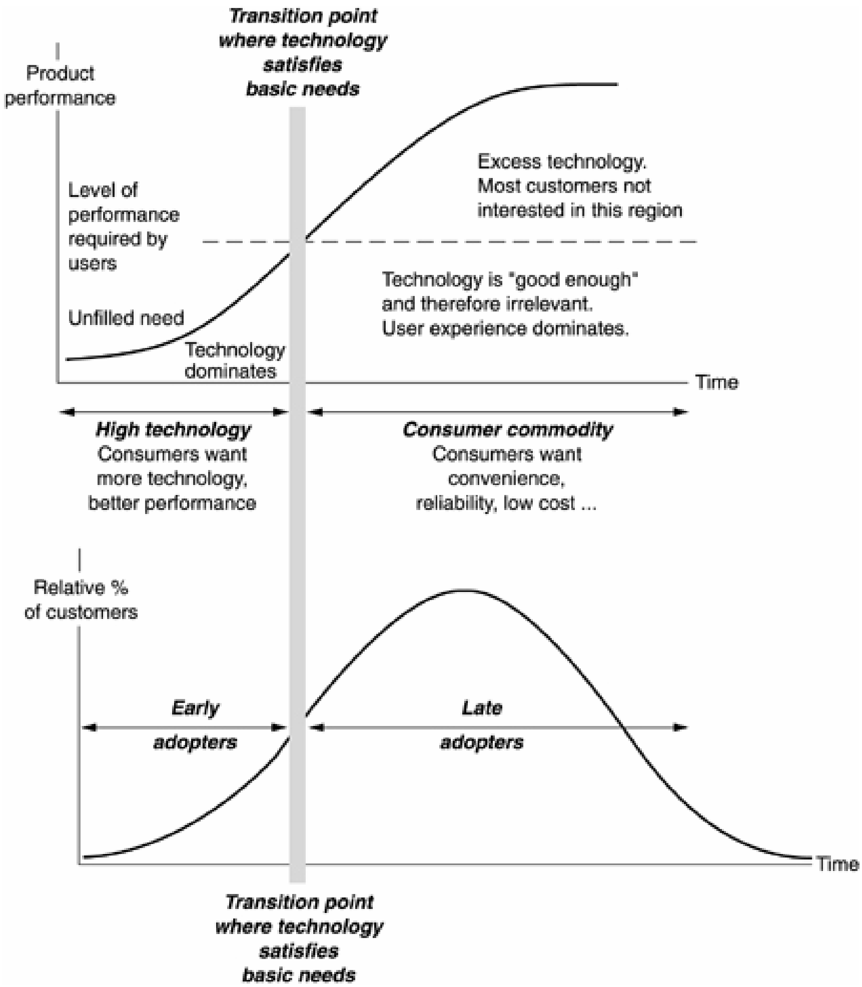


Fig. 1. The importance of user experience varies as a function of market maturity

Vendors of ERP solutions are currently highly profitable, selling solutions that are half as usable as consumer software, and an order of magnitude less usable than consumer websites [10]. They succeed despite the relatively poor UX of their software because no ERP vendor has ever attempted to compete solely based on UX merits to date. This is why ERP companies invest so little in HCI and related innovation efforts. Corporate IT, the customers of ERP systems, generally don't value UX factors that impact end-users today. Steve Jobs [11] recently noted the difference between the consumer and enterprise market dynamics in a recent interview:

"We want to make better products than them. What I love about the marketplace is that we do our products, we tell people about them, and if they like them, we get to come to work tomorrow. It's not like that in enterprise… the people who make those decisions are sometimes confused"

Might this change soon? It might. In general, the population at large is gaining an appreciation for high quality design through the exposure to more consumer products and is starting to demand it from products they use to do work. Another promising sign is that the venture capital community is starting to recognize the value of user-centered design [12].

As a profession, we need to help accelerate this trend by educating those who are influential to recognize that UX best practices can actually help them improve their profit margins in multiple ways. For example, what were the main factors contributing to all those failed IT projects analyzed by The Standish Group over the past 15 years in its aptly named Chaos reports? *Lack of user input, incomplete requirements, and poor specifications.* Those same things are also the root cause of ERP quality issues and the lack of innovation in the ERP market today. The entire ERP ecosystem shares a common viewpoint, for better or worse.

2.3 The Resources Problem— Effect Not a Cause

Values and processes determine resources in companies. As HCI professionals, do you lack the resources to effectively engage users in iterative design? Do you lack the time to develop detailed requirements and specifications? Is your staff unable to find time to collaborate on innovation projects? Contrast the findings of the Standish Group with this quote from Tim Cook, COO of Apple [13] about how his team managed to grow Apple into the 2nd largest public company on earth with a market capitalization of $300 Billion as of 2011, saving it from bankruptcy:

"We're constantly focusing on innovating. We believe in the simple, not the complex. We believe that we need to own and control the primary technologies behind the products that we make, and participate only in markets where we can make a significant contribution. We believe in saying no to thousands of projects, so that we can really focus on the few that are truly important and meaningful to us. We believe in deep collaboration and cross-pollinization of our groups, which allows us to innovate in a way others cannot. And frankly, we don't settle for anything less than excellence in every group in the company, and we have the self-honesty to admit where we're wrong, and the courage to change."

Consider the resource allocation process at Google and 3M. Both are widely recognized as highly innovative companies with strong financials. Both encourage staff to spend 20% of their work hours on self-selected innovation projects [14]. How do these companies manage to find the resources to do this? They recognize that they can't afford not to. To do otherwise would result in a failure to grow their businesses.

In reality, most resource problems have the same root cause. Corporate values and processes that emphasize action over outcome and a focus on short-term costs, versus the long-term returns of design work.

2.4 Good Processes Lead to Effective Resource Allocation and Innovation

Most leaders realize that new, truly disruptive innovations are rare, and far and few between. However, that does not mean they can be left to happen by accident. Big breakthroughs are the typically the result of a long process of extensive exploration to refine the initial idea. One of the greatest technological achievements of the past century, the development of the atomic bomb was the product of a carefully constructed and executed plan that spanned many years [15]. The Manhattan Project showed innovation could be managed into existence. The management lessons learned on that project helped spawn the modern technology revolution.

Real disruptive innovation is rarely an accident, but rather the result of a deliberate application of discovery mechanisms based on the scientific method [16] where organizations focus intensely for years on refining an idea. This requires long term thinking that most ERP vendors lack as they, like many other technology companies suffer from myopic planning cycles that rarely span more than a single year. One could argue that technology makes such long term planning irrelevant, but again there are clear examples of this long-term focus paying off. The consumer electronics hit of 2010? The Apple iPad, a product that has similar characteristics to the Newton MessagePad product it started work on in 1987, nearly 23 years earlier. Recently Apple acknowledged that the 2010 iPad model was over 3 years in the making [17] and its other recent hit, the iPhone, was actually a spin off project of its tablet efforts.

How does one identify and refine new disruptive innovations in a systematic way? There are many well-documented best practices, but these remain poorly integrated into operational processes of today's ERP vendors that focus on product technology development rather than product discovery. Examples include:

1. Lead user studies as defined by Von Hippel at MIT [18]
2. Corporate and user community based prediction markets [19]
3. Ethnographic design methods such as Holtzblatt & Beyer's Contextual Design [20]
4. Applied social network analysis of markets [21]
5. Steven Blanks' Customer Development Model [22]

As technologies continue to mature, most project risks are unrelated to technology factors. Instead, project risks tend to be more about market identification, segmentation, and product design factors. Facebook, for example, did not need to innovate technically, at least for its initial success. Facebook and other web services rely largely on proven technology already in the public domain.

2.5 The Relation to Operational Execution

Bill Gates argues *"that in order for existing companies to be innovative, the key is in execution…"* [23]. Most managers forget design culture, innovation, and operational execution are symbiotic. Even the best ideas will fail to reach the market to become innovations if operational details go neglected due to organizational chaos. In addition, chaotic environments lead to detached staff, and make collaboration with partners and customers difficult. Thus poisoning the well from which innovation springs.

Best practices include defining user experience (UX) related quality metrics aligned with the operational metrics of all functional areas so they are an integrated, instead of an ignored, part of the process. Consider the following best practices:

1. Including UX metrics as part of overall company quality goals and staffing accordingly
2. Defining customer satisfaction goals for all external touch points and linking them to business operations
3. Developing project plans that emphasize learning through design research and iteration
4. Making project failures and lessons learned accessible and part of institutional knowledge
5. Developing budgets considering the true costs of not building the right product or not building the product right.

3 Conclusions

Managers overseeing existing products at ERP vendors tend to focus on short-term cost and risk reduction. They often fail to consider the often high costs of inefficient design processes and the risk associated with building the wrong product (false leads in pursuit of disruptive innovation) by focusing on product development at the expense of customer development and product design.

ERP vendors are at risk of losing market share to competitors who may "out innovate them" by applying UX methods to create superior offerings via refining innovation that may even lower costs or increase profits long-term. Current trends in SaaS delivered ERP are a good example of this. SaaS-based ERP vendors have the advantage of improved feedback loops due to their sales and distribution models that will likely improve their rate of innovation and ability to capture market share.

Both fates are avoidable using proven techniques from other more mature industries such as consumer-packaged goods. Perhaps one of the existing vendors can manage the difficult transformation into a design-oriented culture. A culture that recognizes that technology leadership is no longer sufficient in a mature market. That vendor might survive. However, it is even more likely that a new generation of firms with different DNA will succeed as the market transitions into a phase that values UX. This is more likely, as Christensen, Norman, and others suggest from their analysis of other markets. One only need look to the recent changes in the mobile phone industry to see how plausible such a scenario is.

References

1. Flesch, R.: Why Johnny Can't Read—And What You Can Do About It Harper Paperbacks, New York (1986)
2. Martin, R.: The Design of Business: Why Design Thinking Is The Next Competitive Advantage. Harvard Business School Publishing, Boston (2009)
3. Malone, M.: The Future Arrived Yesterday: The Rise of the Protean Corporation and What It Means for You. Crown Business, New York (2009)
4. Kleiner, A.: The Thought Leader Interview: Tim Brown. Strategy + Business (56). Autumn, New York (2009)
5. Buxton, B.: Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann Publishers, San Francisco (2007)
6. Lafley, A.G., Charan, R.: The Game-Changer: How You Can Drive Revenue and Profit Growth with Innovation. Crown Business, New York (2008)
7. Christensen, C.M.: The Innovators Dilemma: When New Technologies Cause Great Firms to Fail. Harvard Business School Press, Boston (1997)
8. Norman, D.: The Invisible Computer: Why Good Products Fail, The Personal Computer Is So Complex, and Information Appliances Are The Solution. MIT Press, Cambridge (2008)
9. CHAOS Technical Report. The Standish Group International (1995)
10. Sauros, J.: `http://www.measuringusability.com/problem-frequency.php` (2010)
11. Jobs, S.: `http://www.engadget.com/2010/06/01/steve-jobs-live-from-d8/` (2010)
12. Cleveland, B.: `http://www.interwest.com/software-as-a-service/marketing/creating-the-killer-app-starts-with-killer-ueui/` (2010)
13. Cook, T.: Transcript of Apple Q1 2009 Earnings Call (2009)
14. Berkun, S.: The Myths of Innovation. O'Reilly Media, Sebastopol (2007)
15. Zachary, G.P.: Endless Frontier: Vannevar Bush, Engineer of the American Century, New York (1997)
16. Hargadon, A.: How Breakthroughs Happen: The Surprising Truth About How Companies Innovate. Harvard Business School Press, Boston (2003)
17. Stone, B., Burrows, P.: Opening Remarks: The Essence of Apple. In: Bloomberg Businessweek, New York, January 30 (2010)
18. Hippel, E.: The Sources of Innovation. Oxford University Press, Oxford (1988)
19. Taylor, C.: How Bets Among Employees Can Guide a Company's Future (December 9, 2010), `http://www.technologyreview.com/business/26800/`
20. Beyer, H., Holtzblatt, K.: Contextual Design: Defining Customer-Centered Systems. Morgan Kaufmann Publishers Inc., San Francisco (1998)
21. Wasserman & Faust.: Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge (1995)
22. Blank, S.: The Four Steps To The Ephiphany: Successful Strategies for Products that Win. Quad Graphics, San Francisco (2007)
23. Gates, B.: The Wall Street Journal, 1–2 (November 28, 2010)

Future Centered Design: Designing for Sustainable Business

Janaki Kumar and Peter Graf

3410 Hillview Ave., Palo Alto, CA, 94304, USA
{janaki.kumar,peter.graf}@sap.com

Abstract. This paper outlines the changing attitudes of business leaders towards sustainability and the opportunities this presents to product and information designers. Smart business leaders are beginning to understand the importance of sustainable practices to their success. Resource scarcity, increased regulation, and need to safeguard their brand reputation are a few of the key drivers for this mind shift. Just as businesses have to reconsider the way they do business, designers have to re-think aspects of the way they design to participate fully in the creation of a sustainable world. We have to be prepared to rethink our design process, our materials, our product lifecycle, and our notion of customer experience. Information designers and product designers have to work together to create the next generation of sustainability products. While product designers play a direct role, information designers play just an important role by gathering critical information across the enterprise and presenting them in a consumable way to allow decision makers to make informed sustainable choices.

Keywords: sustainability, design, product design, interaction design, information design, business software, user experience design, product design, sustainable design, SAP.

1 Introduction

Sustainability is meeting the needs of the present without compromising the ability of future generations to meet their own needs. - Brundtland, 1987

Over the last decade there has been a significant shift in the way businesses view sustainability. In the past, business leaders may have taken a narrow view that equates sustainability to the environmental movement, dismissing it as irrelevant to their business. Today, business leaders can no longer afford to do so.

Sustainability is now based on the understanding that economies and companies do not operate in a vacuum, but are tightly embedded into societies and the environment. These new stakeholders are forcing business leaders to evolve their perspective on sustainability. Through responsible business practices and sustainable product offerings, business leaders can balance their short term and long term profitability, while considering the economic, environmental, and social impact of their activity.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 449–457, 2011.
© Springer-Verlag Berlin Heidelberg 2011

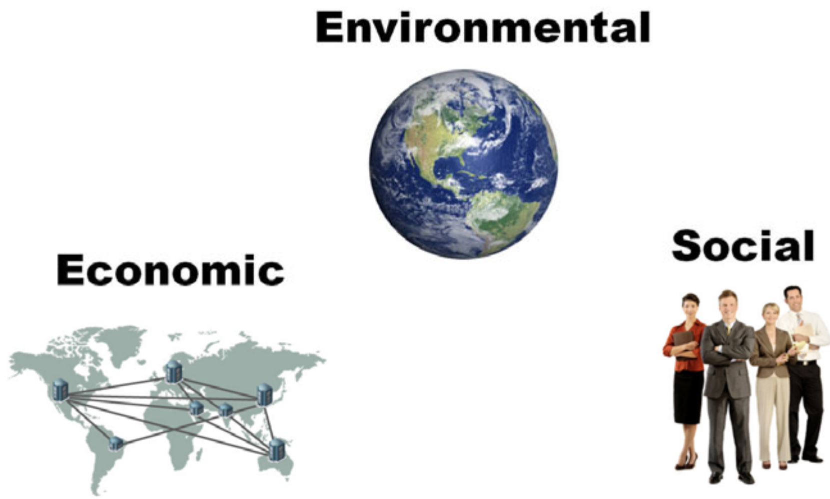


Fig. 1. Environmental, economic and social impact of products

2 Key Motivators for Mind Shift

There have been three main reasons that have accelerated this mind shift.

2.1 Regulation Proliferation

Businesses are faced with an ever increasing number of new environmental, social, and product regulations per year. Such regulations pertain to product and material compliance, employee health and safety, sustainability reporting, and carbon trading. These regulations provide a clear incentive for businesses to think about their operations from a sustainability perspective to minimize or eliminate the risk of non-compliance.

2.2 Resource Scarcity

There is a growing awareness that resources such as oil, water, and raw materials are not unlimited. The current rate of growth and ways of doing business will deplete our planet's resources and threaten our very existence, unless we make serious changes. This impacts businesses directly via the rising cost and extreme price volatility of energy and other natural resources. Resource intensive businesses need to innovate and find other sustainable alternatives or risk going out of business.

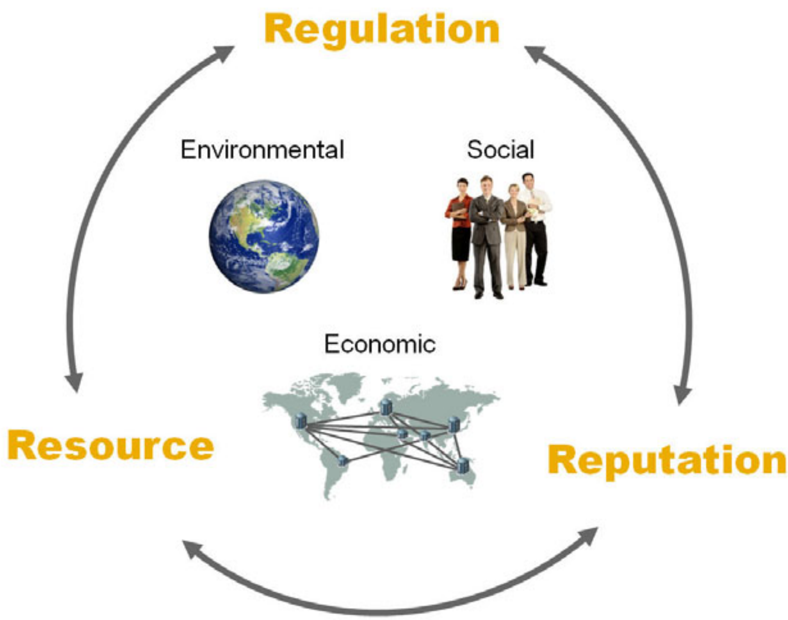


Fig. 2. Key drivers for this mind shift

2.3 Reputation Management

The world is more connected than ever, and as such, there are no longer any local events in the new global economy. With business supply chains spanning the earth in most cases, an ill-advised business decision on one link impacts the entire value chain. Child labor in a major shoe manufacturer's network, and toxic substances in toys comes to mind. Consumers are becoming increasingly interested in the health, social, and environmental impact of their purchasing decisions and are willing to share their experience online with their virtual communities. Businesses no longer have sole control of their brand and are only one of many sources of news about the company. Bad news is impossible to contain and has a severe effect on reputation, sales, and share prices. The recent oil spill in the US is an example of this phenomenon.

3 Business Opportunities for Sustainability

This mind shift translates to real business opportunities. Sustainability is top of mind for many global executives. In a study by the Economist magazine, CEOs see the need for information systems to create transparency around their sustainability Key Performance Indicators (KPIs) and communicate these to their stakeholders. This is closely followed by CEOs wanting to improve the environmental footprint of existing products by using recycled materials, reducing packaging and waste.

To truly seize these opportunities and make a positive contribution, businesses have to rethink some aspects of the way they do business.

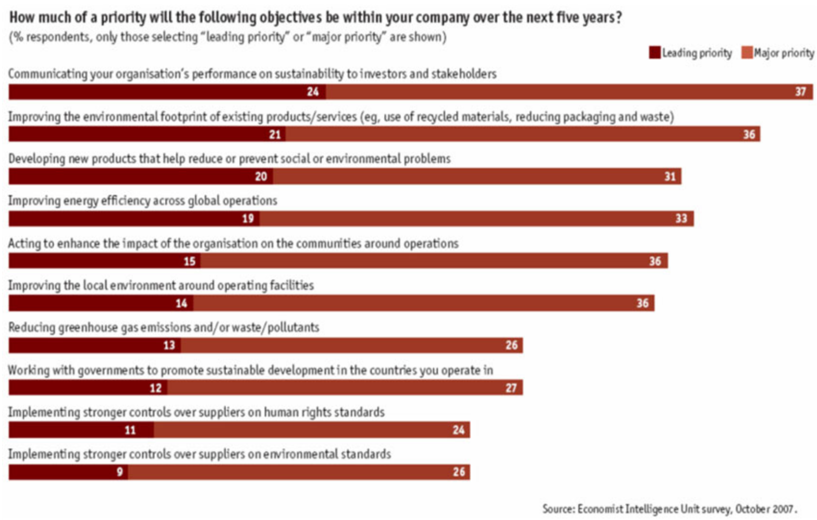


Fig. 3. Sustainability survey of executives

These business opportunities present unprecedented design opportunities for UX professionals. Industrial designers can lead the creation of sustainable product offerings, while HCI professionals such as information designers, information architects, and enterprise software designers can play a leadership role in the creation of information systems that allow executives to monitor and reach their sustainability goals.

Just as businesses have to reconsider the way they do business, designers have to re-think aspects of the way they design to participate fully in the creation of a sustainable world.

4 Re-design Design

Design for sustainable consumption – Bill McDonough in Cradle to Cradle

Future product designers need to take a leadership position in articulating the social and environmental cost of their designs and help stakeholders make sustainable choices. Information designers need to create systems that help gather relevant information across the organization to enable business leaders to make sustainable decisions.

Here are some specific areas of focus to help designers design the next generation of sustainability products and information systems.

4.1 Rethink Design Process

To design the next generation of sustainability products and information systems, UX professionals need to design for the future, not just the user. Design's traditional anthro-centric methods are necessary but not sufficient for the task at hand. Industrial designers need to reconsider user 'needs' when designing a new object and question if these needs can be met in a more sustainable way, either through a different type of product or mode of use. For example, car share systems reconfigure the perception that the consumers need a car when what they really need is transportation.

Information designers cannot merely study current processes in an enterprise when designing the next generation of sustainability software. They instead envision a new kind of enterprise that monitors its sustainability KPIs, similar to its financial KPIs, and creates a reliable and credible report of its sustainability activities. Information systems need to work hand in hand with product design to address issues such as product safety. Today, there is insufficient information available to companies and consumers to make sustainable choices, even if they are motivated to do so.

In many cases, "a typical user" of sustainability information systems does not exist. All organizations have Chief Financial Officers, but it is rare to meet a Chief Sustainability Officer[1]. When the user does not exist, whom does the designer observe and create for? Designers need to go beyond user-centered design methodologies to create sustainable systems for the future enterprise.

4.2 Rethink Materials

Consumers are increasingly aware of the safety of materials. In 2007 alone, 231 recalls in the US resulted in 45,000,000 children's product units being pulled off the store shelves. Beyond putting millions of children at risk, such dangerous products that result in recalls have a negative impact on the company's brand and ultimately end up polluting our landfills having wasted raw materials and energy in its production.

Packaging presents a huge opportunity for CO2 reduction and cost savings for companies. Consumers are questioning the ridiculous amount of packaging they receive when they purchase products. Innovative companies like REI are looking at alternative approaches such as offering customers free shipping if they pick up their online purchases from an REI store close-by, doing away with individual shipping boxes and additional packaging. Lexmark, an SAP customer, set a goal to reduce their packaging by 50%, and in working toward this goal, found that they reduced their energy consumption by 50% as well.

In some cases, the entire industry undergoes a transformation. Digital photography, for example, revolutionized photography and made dark rooms and toxic chemicals a thing of the past. Furthermore digital media is the ultimate sustainable media since they are reusable by design.

---

[1] SAP is an exception in this respect, since Dr. Peter Graf, the co-author of this paper, is the Corporate Sustainability Officer of SAP.

Fig. 4. 2007 Year of Recall - kids in danger

4.3 Rethink Product Lifecycle

Traditional design is concerned with the creation and consumption of products. Sustainable design needs to go beyond this by taking into account raw materials extraction, materials processing, component manufacturing, assembly and packaging, distribution and purchasing, installation and use, service upgrading and maintenance, and disposal and recycling. Thinking through each aspect of the lifecycle, as outlined in William McDonough's book Cradle to Cradle, highlights the designer's responsibility to create sustainable products.

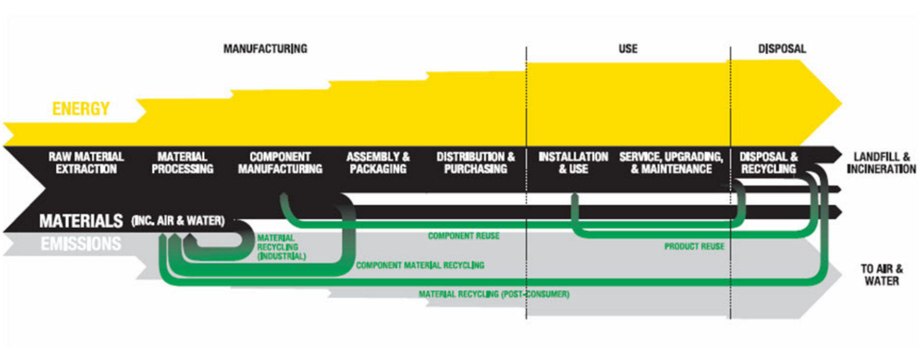


Fig. 5. Total Life Cycle Analysis of products

Today's information systems provide insufficient and incomplete information to decision makers to help them understand the sustainability impact of their choices. Again, information designers and product designers have the opportunity to work together to create the sustainable future.

4.4 Rethink Consumer Experience

Today, the financial impact of the product is made visible via the price. How can designers make the consumer aware of the environmental and societal impact of their choices?

There are some leading companies like Patagonia that have made the product footprint information available on their website.

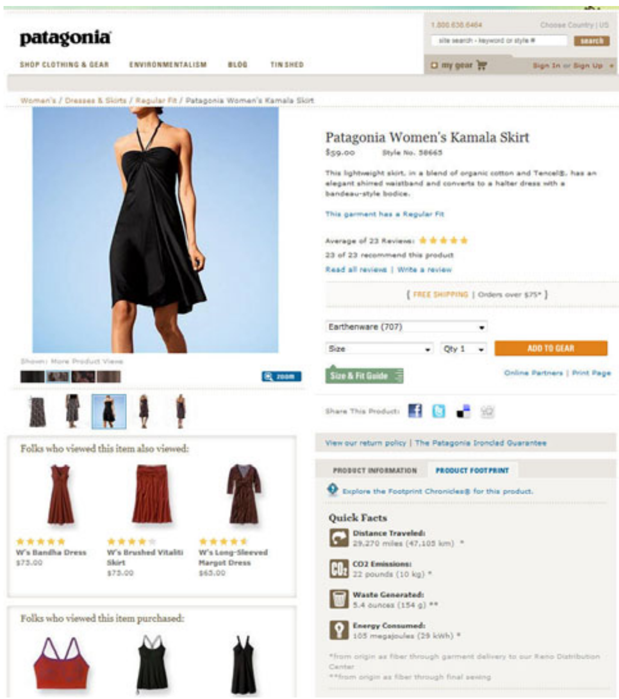


Fig. 6. Quick product footprint information in addition to price and color

Organizations such as Carbon Trust have put in place ways for manufacturers to get a carbon reduction goal and make this information available to consumers.

Here is a vision for mainstream manufactures like Pepsi calculating and displaying their Carbon footprint at the manufacturing, usage, and recycling phases of the product's life cycle.

Fig. 7. Detailed product footprint information

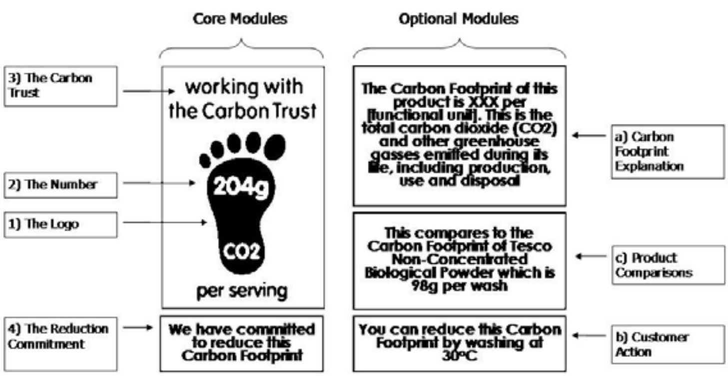


Fig. 8. Carbon Trust logo with explanation

There is a lot yet to be done in this area. Information designers need to create systems to gather product footprint information across the supply chain, while product designers need to present the summarized information to consumers.

Traditionally, design is evaluated on its usability and emotional qualities. We believe that sustainability will become a future product quality that consumers will consider. Savvy consumers are willing to pay extra for the sustainability qualities they value, such as product safety and low product footprint. Therefore, product and information designers need to work together to understand what motivates consumers and help them make sustainable choices.

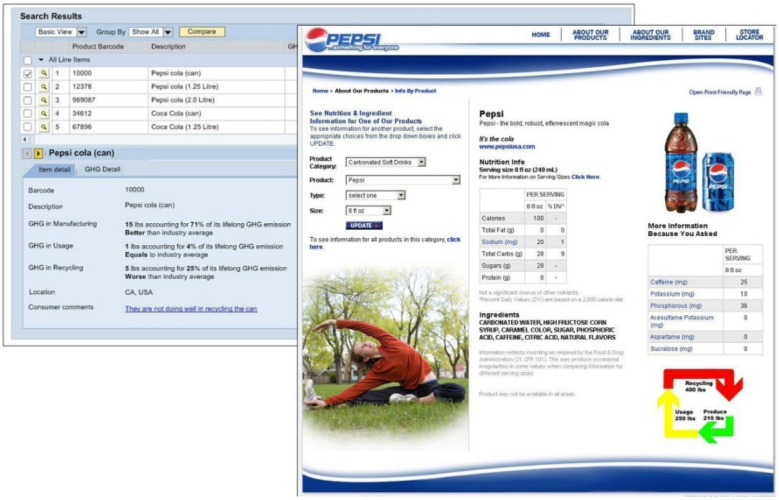


Fig. 9. Product footprint for a can of Pepsi – a vision from SAP

5 Conclusion

There is no better time than the present to focus on sustainable design. Businesses are becoming more aware of the financial, social and environmental impact of their offerings, encouraged by increased regulations, consumer awareness, and scarcity of resources.

The growing middle class in Asia, aspiring to a western lifestyle, are poised to create unprecedented demand for products and services. Our current unsustainable practices coupled with this raising demand could greatly exacerbate our problems, if we do not address them with a sense of urgency.

As any thoughtful designer knows, design problems are merely opportunities for creative design solutions. Therefore, there is no doubt that designers will rise up to the challenge of creating a sustainable world.

The principles of sustainable design are compatible with the principles of any good design philosophy. Design for sustainability places additional emphasis on reducing the environmental impact of a product during its manufacture, use and disposal, or reuse. It uses strategies such as avoiding use of toxic substances during production, minimizing materials used, minimizing energy or water required during use, and designing for repair, reuse or disassembly and recycling.

Information designers and product designers have to work together to create the next generation of sustainability products. HCI designers play just an important role in gathering critical information across the enterprise and presenting them in a meaningful way to decision makers to enable them to make smart, sustainable, business decisions.

References

McDonough, W.: Cradle to Cradle

Schweikhardt, E.: User Centered Is Off Center. Interactions, May/June 2009, pp. 12-15 (2009)

Enterprise UX Design in China

Dongyuan Li

Huawei Industrial Base, Bantian Longgang, Shenzhen 518129, P.R. China
ldy@huawei.com

Abstract. In China, there are many large ITC enterprises and many roles in these enterprises, but UX of working software are not good because many products are developed without UCD method for many reasons. The author gives a way to simplify users by personas through user research and concludes some rules to make design more satisfied by Chinese users because Chinese special culture and its rapid growth, all workers want to improve their skills and ability in order to work more efficiently and find more opportunities for better future.

Keywords: 5E[1], Enterprise user, UX, UCD, UE.

1 Introduction

In China, there are many large ITC enterprises such as China Mobile, Huawei and etc. There are many roles in these enterprises. But UX to working software are often experienced common or hard to use because many products are developed without UCD method. There are many reasons for bad usability: most enterprises do not aware that UCD is a good method to improve software usability and user experience, few requirements are given for these products and no professional experience related engineer is involved in developing process; even though some large OA platform are designed in good architecture, they do not put UCD work in customized development, for many managers think employees should work hard, but not play or waste time on pleasures feeling when using software, and the so-developed software bring many complaints and even make work efficient lower than before; in some enterprises, they have UCD team but focusing on products outside because customers' values are more important and can be embodied on products.

The author has studied many interesting users in enterprise, and find Chinese society have now entered a rapid developing period and tens of thousands of enterprises grow from small to large. Roles in enterprise are many and users in these enterprises often want to know more and get more skills to have the abilities to do something successful in the future. Many users show extreme interest in extra functions that often misleads the appraisals so that they do not give correct analyze to these feedback and do not realize software enough flexible and usability and make a common software in the market for enterprise users.

In this paper, the author want to give some special findings in China, and give explanation through culture factors such as power, Doctrine of the Mean and its

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 458–464, 2011.
© Springer-Verlag Berlin Heidelberg 2011

hieroglyph language, and suggest usability engineers in user research in China, motives, tasks and goals should be deeply mined in order to understand user requirements.

The author also categorizes most users into three roles: manager, technician and operator. So this can simplify user personas to design software and help developers understand their users quickly and gives special design rules in China to help designers to create better interaction.

The paper is organized as the following. Section 2 reviews relevant literature. Section 3 introduces research findings, and Section 4 presents solution and the suggested design rules and in Section 5, the paper discusses some thoughts with future work directions by the author.

2 Related Work

Most UX study and design are concentrating on end customers or consumers such as Mobile phone, website design, social media, game, and so on. There are few studies on Chinese enterprise users, especially analysis in user roles and design rules.

There is one paper on web: "OA System Page Design by Core Idea of UCD" [2]. This paper mainly show how to use UCD method in an enterprise OA system development and give style change to realize users' satisfaction. The only one design case described that the original style confirm by managers was not acceptable by employees and I will explain why this often occurs.

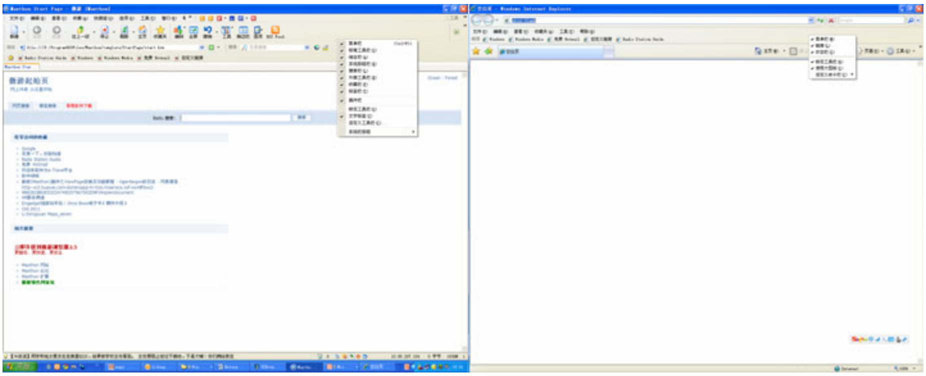
3 Research Findings

In China we find many cultural factors affect user research quality such as power, Doctrine of the Mean [5] and its hieroglyph [3] [4] language, and also rapid developing environment in China effect users' recognition to new products. So in UX research, motives, tasks and goals should be deeply mined in order to understand user requirements.

In an ITC enterprise most users can be categorize into three roles: manager, technician and operator. Management users (later we call this role as manager) understand, guide and supervise team member or interface to realize enterprise goals such as CEOs, CIOs, CFOs, and other officers and managers; technical users realize their technical work value through IT tools such as IT supporters, technicians, R&D developers; Operator users operate terminals to serve customers well and solve their problem efficiently.

3.1 Chinese Characters Make Read Simplier

Chinese users need more meanings in one page. This result, for many reasons, cannot get from a strictly study between comparison. But I met some researchers, they told me in western users' eye, a good design page directly translated to Chinese language often feel too simply and no interested in. When I design concept model, I found many operators do not think more complicated are unacceptable, they expressed wills to learn if they think this product design will help them to get better efficiency and finish more tasks. This do not mean complicated data push down to user is better than that in western society, it means more work will be done to architect the framework and balance the elements in design. You can find many technical users like to use Maxthon Browser than IE or others as follows.



The Maxthon Browser has more function and more good information, tips and selection for them.

The reasons that make this result maybe more complicated are that Chinese characters are belonging to Xingyi text which means it express its meaning through its own form with text.

Fig. 1. Chinese character evolution from Carapace-Bone-Script to Modern Writings

Table 1. Chinese words have more relations than English words

Chinese Word	E s W
	ox
	bull
	cow
	calf
	milk

3.2 Doctrine of the Mean Make Analyse and Design More Complicated

The Doctrine of the Mean has deep effect into many Chinese behaviors. When you ask ideas of Chinese users, they will not give bad answer because they do not want their guest to be disappointed. In product design evaluation, many engineers may get good or not bad appraisal, but when product is online and available, the market does not seem like the expectation and a lot of questions and problems feedback to developers.

So, user research in China needs more professional researcher or usability engineer to do. When we facilitate a study or an appraisal, facilitator needs to prepare right questions and communicate more deeply into user's heart to understand his motives, tasks, goals and their relation between them. So these relations analyze make result of study truly and authentic.

There's another reason I think may help why this phenomena often appear: Chinese society grow rapidly and people want to get good relation and do not want to waste time on unfamiliar things.

This can be proved from many cases. I just give two examples, one result is good and another is unsatisfied to many users. The good one is that they (usability engineers, later abbr. as UE) spent a lot time on selecting typical users and categorizing them into fresh, common and skilled users. UE studied these users and analyze the interacting habits between each group and UE design concept and low fidelity prototype design and evaluate carefully, even using eye-tracking to analyze different design details, and give UI specification to make developers understand how to design each component. Finally the software was successful and good appraisal by enterprise users. Another case UE mainly focused on problems when users use the old one, and did not analyze user group and their scenarios and tasks into enough depth. When UE finish low fidelity prototype design evaluation, high appraisal seems the design maybe successful. But before and after the product released, many unfavorable opinions fed back to the development team. When we traced back, we found there are many reasons might affect design quality, but for UE, should analyze the internal thinking by users such as motives, goals and logic relation between these. If UE only depend on the simple answer by user, many helpful ideas will lose.

3.3 Power and Non-power Influence

After many researches in enterprises, we find user investigation would meet with many difficulties without support by users' manager. For enterprise user has his or her own work and user research needs user spend enough time on study. It is best to invite the manager to give an open support and arrange team members' plan. But most difficult thing is when user is a manager, how to make user research? And higher the position is, more difficult to study the manager. The best way is to get support high level manager's support and propaganda on bulletin board. But much higher manager is, much busier will be. So observation often is used.

Senior expert often has great effect to other users, and they will give no good result if expert sits beside them. For Chinese user, they often think senior and expert have deep thoughts and they should follow the right thoughts. But often these experts want to show their abilities and will say that "all is ok, it depends on your model". And UE cannot get more information about user or prototype. In this circumstance, 1to1 interview will be more effective.

In an ITC enterprise, there are more tens of roles when we do user research in a complicated workflow or OA or other such computer-software system. There are different levels of managers, different skills of technicians, various services of customer representatives, and they have many cross points.

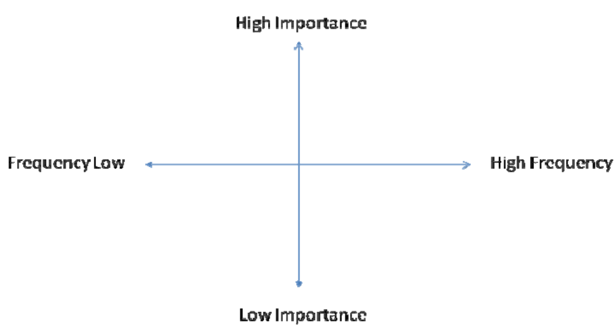


Fig. 2. Using Importance and Frequency to analyze

We use the following two dimensions to analyze user's task and we find managers seldom use Enterprise IT equipments but they bring product requirements and decide which product should be bought. Operator seldom appears in buying, but they require their task should be quick finished by using computer. Technicians' tasks are too many and information architect in system and interaction design should be design flexibly to support them to finish different tasks in different circumstances in order to ensure enterprise run safely, stably and effectively in each sector and each level.

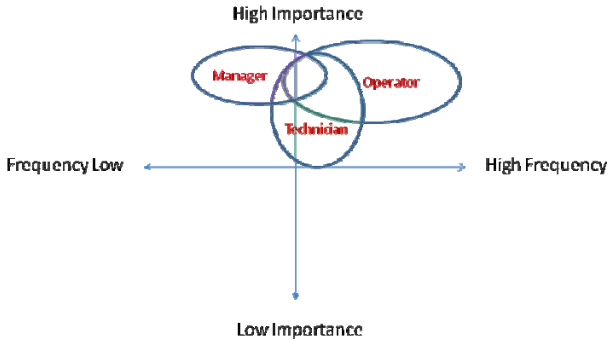


Fig. 3. Manager, Technician & Operator's relation with working products

Though modern Hi-tech make an ITC enterprise organization very complicated, but most users can be categorize into three roles: manager, technician and operator. This chart above may simply user personas, but not shows that user research will not be needed in enterprise. We should think how to consider each cognitive way and design balance between these factors. We can see from this chart that manager often fill importance with team member, but he or her will not use computer as often as team members for the working software. The difference between their behaviors will make different feel or requirement for the product. For example, a manager may want to see the color more brilliant and attractive but an operator may hope to work with soft, gentle colors.

4 Design Rules

Nielsen gave 10 usability rules and Whitey Quesenbery also gave rules for 5E. I give special design rules in China to help designers to create better interaction.

1. Give users enough selections. For Manager, technicians and operators have different task and if we cannot give enough selections such as customization or settings, you will find they requirements could be hardly satisfied. Rapid developing society makes enterprise users their abilities grow quickly and they often want to have more facilities in one window.
2. Information Architect should be carefully used and many design techniques such as navigation, metaphor, tips and notes make task easy to learn and finish. In China, user want to work quickly and they working responsibility are often changing. If a product in Information, Interface and Interaction are not integrated to good design in finishing task or frequently operations, it will bore users and it will be more difficult to collect data and update.
3. Give draft or small space for recording user's previous information. For Chinese users, they often meet different or temporary change for the present task. If task is too long or their working tasks are too many, a user often has to or is apt to change to another topic.

4. Clear working status data. This will help managers to understand enterprise's working status, and operators know how much their task has been done and technician will know tasks those are assigned and importance between them.
5. Color should be designed carefully correspondently to the working environment. Managers often want colors to be brilliant and attractive, but operators want colors to be comfortable and gentle.
6. Give feedback channel online that shorten the distance between user and maintenance supporters.

5 Conclusion and Future thoughts

In this paper, the author analyzes Chinese culture effects and proposed scientific and professional way on user research. The author also gives three user categories and recommendation rules for design and balance in enterprise software in China.

The research finds may lack scientific proofs and need our practitioner to do more work in this domain. And enterprise products have many types, and we only have much experience in OA, communication and working software (CRM, BSS, BI and etc.). In our team, we accumulate personas database and design framework and components, in order to improve design efficiency and quality. We plan to do this work in new service area.

References

1. Quesenbery, W.: Using the 5Es to Understand Users. The 5E means five dimensions of usability: Efficient, Effective, Engaging, Easy to learn, Error tolerant
2. Wang, K.: OA System Page Design by Core Idea of UCD. Magazine of Modern Enterprise Culture (published on 11th 2008)
3. Chinese Character: an explanation from the web dictionary lib: `http://baike.baidu.com/view/1712.htm`
4. Chinese Character: Wikipedia: `http://zh.wikipedia.org/zh/%E6%B1%89%E5%AD%97`
5. Doctrine of the Mean: Wikipedia: `http://en.wikipedia.org/wiki/Doctrine_of_the_Mean`

Changing the TV Industry through User Experience Design

Daria Loi

Intel Labs, Intel Corporation, 2111 NE 25th Avenue, Hillsboro OR 97124, U.S.A.
daria.a.loi@intel.com

Abstract. This paper discusses a process used to enable the TV industry toward major changes around the consumption, navigation and sharing of content. A chronographic description of key milestones – from ethnographic research to concept design and customer impact – is provided. The illustrated case study is grounded in the context of a large silicon manufacturer, where notions of change-through-design, innovation and customer impact have distinct connotations and set boundaries, especially when compared with medium sized firms, original equipment manufacturers or academic contexts.

Keywords: Design, Ethnography, Human Factors, Interface, TV, UX.

1 Introduction: TV Industry and Research

In this section I will introduce the field of study and offer a brief overview of the methods, tools and strategies underpinning the reported case study, which are rooted in a mix of ethnographic, human factors and design thinking.

As a technology that has been around for over 70 years, television continues offering something that people highly value: social content consumption. For the 2009-2010 broadcast seasons Nielsen [1] estimated that the total number of television households within the U.S. (including Alaska and Hawaii) will be 114.9 million. For a country that counts approximately 115 households, this is a remarkable number, which demonstrates the pervasiveness of this *old* technology we call TV. Strategy Analytics [2] forecasts that global digital television subscriptions will grow from 484 million in 2010 to 887 million by 2014 – on a regional basis, Asia Pacific, Central and Latin America and Central/Eastern Europe are expected to grow fastest in the next five years.

While the number of TV subscriptions kept growing, this popular technology failed however in progressing the way other technologies did, leaving the living room with a comfortable void, where connectivity, social networking and contextual information are most likely offered by other devices (for instance: smartphones, netbooks, laptops). This void seems to be reflected in recent estimates that about 800,000 U.S. households cut the cable in the past two years (cumulative number, with 600,000 people cutting the cord in 2009) and that the number is forecast to double to 1.6 million [3]. Although these are small numbers in the context of an $84 billion Cable/Satellite/Telco TV industry, they indicate a shift to web-based TV viewing – a shift that the industry is not ignoring.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 465–474, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Within this context, a number of new technologies and opportunities have been explored by TV industry players. One of them is the opportunity of merging the world of TV and Internet in one platform for the living room. The case study here reported relates to work conducted at Intel Corporation in relation to Intel® architecture-based system-on-a-chip (SoC) media processors including Intel® Atom™ processors for TV [4] and to its smart TV project [5].

To develop its smart TV enabling products, Intel's Digital Home Group focused on a process grounded in user experience research, where ethnographic insights, design enabling activities [6] and human factor competencies helped explore, develop and test novel scenarios and usages; define and drive new capabilities; imagine, develop and test reference designs; and drive new opportunities with relevant stakeholders.

Intel's effort to enable the smart TV experience is about "enabling consumers to interact with their TV like never before, seamlessly integrating a broad array of Internet content, broadcast programming, personal content, and virtually unlimited applications – all viewable on one TV screen" [5]. This experience was developed through careful user experience research and design. The case study reported in the following sections illustrates an example of how this was achieved.

2 Case Study Overview

In this section I provide a succinct summary of the key steps that define the case study here reported. Section 3, 4, 5 and 6 will further elaborate on each of the points here briefly summarized.

In the past 5-6 years researchers, designers, planners and technologists in Intel's Digital Home Group worked to understand the living room experience and develop silicon products capable of enabling a user-centered merging of broadcast and broadband content. Within this context, user experience researchers and designers conducted a number of exploratory studies targeting various aspects of the digital home, including TV experience. As part of these endeavors, in 2007 I co-conducted an ethnographically informed study focused on home-care practices.

The insights on media collection and navigation gathered through this ethnographically informed study was the foundation for an endeavor I and another colleague drove to design a media aggregation widget for TV. During the widget development we discovered limitations in the framework that did not enable us to design an experience that could fully mirror our user insights on media collection practices. In response to this challenge, we therefore decided to develop a concept demo that would enable us to illustrate some media navigation opportunities within 3D user interfaces (UIs). The user testing of this and other 3D UI demos illustrated the possibilities offered by rich interfaces as they seemed to better match people's expectations around what is displayed on the TV screen.

Following the success of these UI experimentations and of a number of user experience studies and tests conducted by my team, I was then asked to drive the development of a *vision demo* for a keynote address at the 2009 edition of Intel® Developer Forum. The positive responses to the demo by customers, press and analysts stimulated a number of follow ups and initiatives, enabling the TV industry toward major changes around the consumption, navigation and sharing of content. The following four sections will elaborate on each of the points I have briefly summarized in this section.

3 The Ethnographic Chapter

As previously mentioned, in 2007 I co-conducted an ethnographically informed study focused on home-care practices. The study, conducted in Sweden, China and Indonesia, included face-to-face interviews, household and contextual observations, expert interviews, the use of Cultural Probes [7] and to some extent Playful Triggers [8,9]. Although the study unfolded a number of matters related to home practices, I will here focus on a specific area, since it was in my view the seed that started the developments at the heart of this paper.

Fig. 1. Cultural Probes [7] in Indonesia (left), China (center) and Sweden (right)

As in most ethnographic studies, once we returned from the field we found ourselves with rich insights and the urgency to cut through the data to provide useful directions to our planning and technological counterparts in the organization. One of the areas that strongly emerged from the data was that of storage practices [10].

When looking at storage practices, an important issue that emerged from the data was that, in everyday contexts, digital storage is a practice rather than a technical device: it is the work of managing content – keeping, protecting and finding the things one cares about. We consequently highlighted that digital storage solutions entail more than simply building better technologies, because *storage* demands to perform tasks consistently and repeatedly. We found that people are very rarely pro-active in their digital storage management practices – they are rightfully so absorbed by their everyday tasks and lives. We consequently recommended that it is better not to rely on people to have pro-active practices and to instead develop proactive storage solutions that demand fewer repetitive and tedious tasks from them in order to keep, protect and find the things they value.

Our insights arrived at a crucial moment – the topic was of great interest to many internal stakeholders and deeply impacted their thinking and development. A good example is the article *Why We Need Whole Home Storage Architecture* [11] where colleagues at Intel articulated – thanks to and through our supporting data, insights and recommendations – the needs for a novel home storage architecture called Home Media Aggregation.

4 TV Widget Development

As mentioned in the previous section, the data and insights that arose from the homecare study and our related storage solution recommendations deeply influenced the thinking around Media Aggregation architecture.

At the same time, two important developments were under way: firstly, the Consumer Electronics (CE) industry was increasing its focus on storage and media access on TV; secondly, Intel was collaborating with Yahoo! Inc. "to provide a full-featured software framework named Widget Channel that allows TV viewers to enjoy rich Internet applications called TV Widgets while watching their favorite programs" [12]. Widget Channel, "designed to run on a variety of connected CE devices", provide a way to "personalize, enjoy and share Internet content and services on TV by enabling multiple Internet applications to be displayed on the TV screen concurrently with video programming" [12].

Considering the above, I and another colleague started working with internal stakeholders to drive the design and development of a media aggregation widget for the Widget Channel. The *MyMedia Widget* aimed at demonstrating "how personal media like music, photos and videos can be incorporated into a TV Widget" [13] and to do so in light of people's practices around media consumption, collection and navigation. Through this widget users can access from their TV photos, music and videos that reside on any connected device, as if content were stored in one location.

This widget was tested and refined through User Experience Assessment (UXA) and launched at the Consumer Electronics Show (CES) in 2009, followed by a number of ad hoc refinements for target customers.

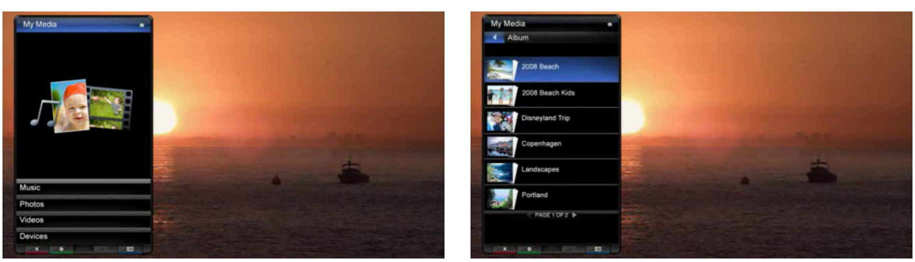


Fig. 2. MyMedia Widget: main menu (left) and submenu (right)

5 Toward 3D User Interfaces

In late 2008 and early 2009 many of the UX assessments we conducted around user interfaces indicated that users' expectations around TV UIs are more sophisticated when compared to their expectations around other devices. At the same time discussions between user experience and technology counterparts in our group directed our interest toward 3D user interfaces for TV platforms as we felt they could offer opportunities in the context of the above-mentioned expectations.

During the development of *MyMedia Widget* (as reported in the previous section) we realized some of the insights we gleaned through our ethnographic studies and UX assessments could not be implemented within the software framework we were operating within – we had restrictions that did not enable us to fully develop all our ideas around TV-based media navigation and consumption.

The mentioned focus on 3D UIs for TV was timely as it enabled us to explore and propose the development of a new widget, one unrestricted by specific software frameworks. *MyMedia 3D* was therefore developed in early 2009, to showcase a number of concepts related to personal media (3D navigation, consumption, sharing, contextual recommendations and promotions). This 3D widget had two states (a side wheel and a full wall) to enable users to navigate their own content in quicker or more extended manners. The TV programming stays in its full screen mode when the side wheel is used, while it becomes a picture-in-picture when the user chooses to enter the media wall mode. In both cases a range of options are offered, linked to each media item (picture, video or music): 'More like this' (to find similar types of media in one's personal collection, on the web or in social networks); 'Recommendations' (to get related media or content that might be of interest to the user); 'Your friends liked' (to view other media recommended my personal network); 'Send to' (to share specific items with others); and a promo banner, to advertise related items of interest (for instance, the possibility to buy a concert ticket for the artist the user is browsing in that moment).

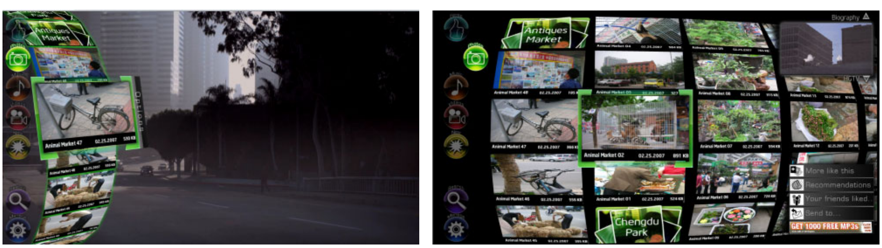


Fig. 3. MyMedia Widget 3D: side wheel (left) and media wall (right)

In March and June 2009 we conducted two UXA studies loops to test and iterate the design of MyMedia 3D. In both sessions, users' response was overall positive and encouraging, with users providing specific emphasis on:

- 3D look – for its *cool factor* and *smoothness*;
- opportunity to visually view many elements at once;
- possibility of sharing media;

- possibility to get related content, especially in relation to music; and
- live TV as picture-in-picture during non-immersive activities such as browsing lists of photos.

In the first UXA round however a small subset of participants expressed they felt overwhelmed with the graphic layout of the media wall and its animations. This prompted us to reiterate and modify the design, with a resulting positive progression in the overall response to 3D interfaces, depicted during the second UXA test as "cool" yet balanced by usefulness and subtler effects.

6 Concept Demo and Reference Design

In mid 2009 I was asked to develop a concept demo to kick off the keynote address of Intel Digital Home's Vice President at Intel Developer Forum (IDF). The aim was to demonstrate – vision demo at hand [14] – a possible future of the TV experience, based on user experience research, definition and design.

The demo, named *TV Reinvented*, was created by developing a cohesive experience capable of showcasing the top value propositions I felt strongly about thanks to our UX research and testing:

1. Browse content and change channels through 3D user interfaces
2. Share/send personal content with/to others
3. Access and receive contextual information and recommendations
4. Make and receive video calls on TV
5. Watch videos together from different locations
6. Gesture-based navigation
7. Voice-based search

The first three items on this list directly derived from the work and experiences described in the previous sections of this paper, strengthened by additional ethnographic research conducted by my colleagues in the previous years. The fourth and fifth items represent strong usages that our team tested a number of times in a number of geographic locations over the years. The last two items derived from research, development and testing I conducted at Intel since 2008 around advanced remote controls with gesture, voice and touch modalities [15] and from personal research on natural interfaces [16].

Fig. 4. Stills from animations developed to test usage concepts of gesture-based navigation for TV (left: camera-based gesture; right: device-based gesture)

Following the overwhelmingly positive responses to the *TV Reinvented* demo by analysts and tech experts [e.g. 17,18,19] and by customers that, inspired by the demo, designed new proof of concepts for their products, we decided to develop reference design code for some components of the demo (channel wheel and media wall). The reference design of *TV Reinvented* was showcased at CES 2010 [20] and aimed at enabling our customers to develop similar experiences within their own software frameworks and products.

Since its 2009 concept version at IDF and its reference design showcased at CES2010, *TV Reinvented* has travelled around the globe and in many initially unforeseen places – besides being shared with many customers, it was showcased at IDF Beijing (PRC), Sponsors of Employee Expo (USA, PRC, Malaysia and UK), WCIT 2010 and IBC 2010 (Netherlands), DTTC 2010 and Comic Con 2010 (USA), and iFA 2010 (Germany); presented to academics and researchers in USA, Europe and Australia; showcased in a number of Intel executive keynotes; and was even inspiration for Intel® Consumer Electronics (CE) initiative advertising.

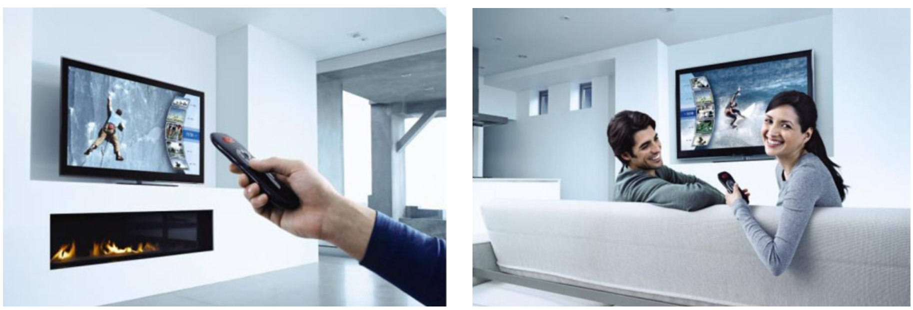


Fig. 5. Promos for Intel® Consumer Electronics http://intelconsumerelectronics.com

7 Observations and Implications

In this final section I will provide some high level observations and implications arising from the case study and discuss the roles that design can, could and should play as a catalyst for change. It is important to emphasize that the illustrated case study is grounded in the context of a large silicon manufacturer, where notions of change-through-design, innovation and customer impact have distinct connotations and set boundaries, especially when compared with medium sized firms, original equipment manufacturers or academic contexts.

Firstly, I would like to point out the roles that user experience research played, since it enabled the development of ideas grounded in everyday life, making conceptual and design propositions strong in their UX intents. This was achieved by carefully considering existing and relevant qualitative and quantitative data; continuously questioning and re-iterating usages and designs; testing concepts through User Experience Assessments; and integrating new pieces of ethnographic and design research each time it was made available.

Secondly, one of the unanticipated outcomes of the case study was that it opened up new opportunities with internal stakeholders asking for design expertise to re-conceptualize – from a UX and design perspectives – their technological frameworks. A good example is Intel® Multi-Application Framework that I was asked to re-design from a UX/UI perspective and that was launched at IDF2010 [21,22]. Beyond the project in itself, this outcome has a great value as it indicated a welcomed growth in the organization's approach toward the value that design expertise brings. Moreover, it opened up the opportunity for close collaboration between design and software engineering teams in US and PRC, which resulted in the development of great trust among all parties and the capability to work together, understand each other's disciplinary language and achieve the best possible outcome, collaboratively. In a way, it could be argued that the project itself played the role of a Playful Trigger [8,9] among design and engineering teams, creating the communication bond necessary for effective collaborative practice to emerge and be sustained.

Another factor that I would like to emphasize is that, while UX insights were being accumulated over a number of interlinked projects and unforeseen opportunities kept emerging, an entrepreneurial creative attitude was there to support this continuously morphing project and its at times serendipitous developments. Decisions were no doubt guided by user experience insights and in some cases by corporate needs and wants yet there is something to be said about the role that what I will here label as *designerly gut feeling approach* played (some might call it *intuition*).

This approach is what made in my view a difference, freeing the team from corporate conventions and enabling them to imagine and create new possibilities instead of feeling somewhat compelled to be safe in the known. I believe that the approach made it safe to play and design through play, creating new scaffolds instead of working within given ones.

Is this something that could be replicated? I believe so, however it comes with a price, since *designerly gut feeling approaches* require a great deal of managerial support, enablement and trust. This is likely to naturally exist in design-led institutions and design consultancies yet it is less likely to be found in large technology companies due to their nature which is typically less about *guts* and more often about *safety*. There are of course grand examples of large companies that put safety a bit on the side to take on board informed dream-decisions – yet the vast majority of the large tech companies are, I believe, paralyzed due to their fear of being different and their impulse to play it safe.

This is why the reported case study is a story important to be told, as it gives an example that things can happen, despite fears and obligations. Again, is this something that can be replicated? I believe so, yet we need managers capable of working with designers and the creative workforce. We need them to be trained so they can understand how *designerly gut feeling approaches* work and learn how to properly support them instead of discourage or deflate them. This is not easy, especially within large tech companies where being a technical contributor and a good manager at the same time is a very hard task indeed.

Regardless of this debate however, what appeared fundamentally and clearly obvious during this case study is that design can play significant change agent roles and that it is a capability that (even if hard to integrate in some organizational paradigms) can make a substantial difference in the end result, especially when it is matched with ethnographically informed research competencies.

To conclude, I would like to point out where the end results of this case study started, regardless of the role that design and research plays – it started from people, their inspirational perspectives and their stories: from a Swedish chest of drawers and the sudden discovery of unanticipated items; from an Indonesian garage, repository of several broken, forgotten and rediscovered water dispensers; and from a filled-with-DVDs Chinese suitcase that constantly fluctuates between chaos and order, depending on who opens it. When looking at end results and when reading glory articles on new tech products is easy to forget where they all likely started: people's experiences.

Acknowledgments. I wish to acknowledge all colleagues at Intel Corporation that contributed in different ways to the reported case study.

References

1. The Nielsen Company: Local Television Market Universe Estimates, `http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/08/2009-2010-dma-ranks.pdf`
2. Strategy Analytics: Global Digital Television Forecast, `http://www.strategyanalytics.com/default.aspx?mod=reportabstractviewer&a0=5437`
3. Convergence Consulting Group: `http://www.convergenceonline.com/reports.php`
4. Intel Corporation: Smart TV, `http://intelconsumerelectronics.com/Smart-TV/`
5. Intel Corp. Consumer Electronics – Smart TV, `http://www.intel.com/inside/smarttv/`
6. Loi, D.: Leading through design enabling: practical use of design at Intel. In: Design Connexity – 8th International Conference of the European Academy of Design, Aberdeen Scotland, April 1-3 (2009)
7. Gaver, B., Dunne, T., Pacenti, E.: Interactions Cultural Probes, vol. 6(1), pp. 21–29. ACM Press, New York (1999)
8. Loi, D.: The Book of Probes, in Lavoretti per Bimbi–Playful Triggers as keys to foster collaborative practices and workspaces where people learn, wonder and play. PhD Thesis, RMIT University, Melbourne, Australia (2005)
9. Loi, D.: Reflective Probes, Primitive Probes and Playful Triggers. EPIC07–Ethnographic Praxis in Industry Conference, Keystone CO, October 3-6 (2007)
10. Loi, D.: Digital home storage beyond devices: a cross cultural perspective on how people keep, protect and find what they value. Keynote, UPA Europe–European Usability Professionals Association Conference, December 4-6, Turin Italy (2008)
11. Intel Corporation: Consumer Electronics – Home Storage Architecture, `http://www.intelconsumerelectronics.com/Consumer-Electronics-3.0/Home-Storage-Architecture.aspx`
12. Intel Corporation. Consumer Electronics – Widget Channel Overview, `http://www.intelconsumerelectronics.com/Consumer-Electronics-3.0/Widget-Channel-Overview.aspx`

13. Intel Corp., http://www.intel.com/pressroom/kits/events/ces2009/CES_FactSheet_Day1.pdf
14. Intel Corp., http://intelstudios.edgesuite.net/idf/2009/sf/keynote/090924_ek/f.htm
15. Intel Corp. Next Gen Remote Controls, http://intelconsumerelectronics.com/Smart-TV/Next-GenRemote-Controls-Video.aspx
16. Loi, D.: TUIs as mediating tools within adaptive educational environments. In: McKay, E. (ed.) Enhancing Learning Through Human Computer Interaction, pp. 178–191. Idea Group, Hershey (2007)
17. Shiels, M.: Future is TV-shaped, says Intel. BBC News, September 25 (2009), http://news.bbc.co.uk/2/hi/8272003.stm
18. Enderle, R.: Intel's Developer Forum: Heroes and Slackers. IT Business Edge, September 24 (2009), http://www.itbusinessedge.com/cm/blogs/enderle/intels-developer-forum-heroes-and-slackers/?cs=36117
19. Laird, J.: How Intel plans to get inside your TV - IDF 2009: Intel's multi-pronged TV tech onslaught explained. September 28 (2009), http://www.techradar.com/news/computing-components/how-intel-plans-to-get-inside-your-tv-638742
20. World News: http://wn.com/Intel_TV_Demo
21. YouTube: http://www.youtube.com/watch?v=_vFj6Vh22q8
22. Hachman, M.: Intel's design work guided by 'experiences', too. PC Mag, September 9 (2010), http://www.pcmag.com/article2/0,2817,2369032,00.asp

Electronic Invoicing in SMEs

Lasse Lumiaho and Jussi Rämänen

Aalto University School of Science, Department of Computer Science and Engineering
{lasse.lumiaho,jussi.ramanen}@tkk.fi

Abstract. E-invoicing has benefits that are quite valuable to any enterprise, for example, cost savings, automated accounting procedures and increased efficiency of the workforce. This paper reports the results of a study that focused on the ease of use, learnability and low usage of e-invoices in SMEs. The study was conducted with 12 Finnish small and medium enterprises (SME). The paper demonstrates use habits related to the e-invoicing systems and presents positive and negative comments from the users. According to the users, the benefits of e-invoicing are on a rudimentary level. On the negative side, the SMEs cannot realize all the benefits of e-invoicing.

Keywords: electronic invoicing, SME, usefulness, usability, Finvoice.

1 Introduction

The widespread use of e-invoicing has been touted to save 200 billion euros worth of costs, reduce CO2 emissions by 3 million tons per year and free resources for more productive work. Additionally, e-invoicing will also reduce fraudulent invoicing by improving VAT controls. The European Commission Expert Group on e-Invoicing [1, 2] highlights that SMEs need cheap and easy-to-use e-invoicing solutions and services that can be accessed and used without specific IT skills. The group also points out that the ease of use is an important business requirement exceeding simple cost savings. [1]

E-invoices are said to have many advantages over paper invoices as e-invoices can be forwarded via banks similarly to payment transactions. The Finnish e-invoice format, Finvoice, provides advantages for both the seller and the payer. Forwarding of e-invoices is reliable and secure, and the format of the invoice is generic enough to suit the requirements of any business sector. The process of handling incoming e-invoices can be fully automated to reduce handling times and the number of errors. Furthermore, forwarding an e-invoice through the bank network enables delivery notices to be sent to the sender, and the banks authenticate all e-invoice senders providing more security for the payers. The Finvoice format also works on multiple systems providing sellers and payers a variety of software options to handle e-invoices. [3]

The study was conducted in collaboration with three Finnish banks. Each bank provided four SME-sized customers for the study. We carried out expert evaluation as well as field studies to test the usability of sending e-invoices in the three Internet bank systems. In addition, we wanted to find out use habits related to the e-invoicing process.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 475–484, 2011.
© Springer-Verlag Berlin Heidelberg 2011

1.1 Research Goals

In our research, we have noticed that the diffusion of e-invoicing has been low in SME-sized companies. The motivation for this study was to investigate the reasons for the low usage of e-invoices in Finnish SMEs and the low utilization rate of opened invoice services. We were particularly interested in creating and sending e-invoices, and therefore, receiving and paying e-invoices were excluded from the study. In the beginning, we contemplated that a troublesome adoption phase, bad usability of the banks' e-banking systems as well as a mismatch between e-invoicing and established ways of creating invoices could be the possible reasons for low usage and the low utilization rate.

1.2 Research Methods

The study was carried out in two phases, namely expert evaluation and field studies. The Internet banking systems and the e-invoicing services are usually put into use without any instructions. Therefore, we would need to carry out expert evaluations that analyze the ease of learning as well as the match between users' needs and the system's functionality. In addition, the expert evaluation method should bring supplementary data, as only a limited set of users participated in the study. Cognitive walkthrough, which was chosen as the expert evaluation method, focuses on ease of learning but also addresses functionality and ease of use as they contribute to the ease of learning [4]. In our reasoning, the cognitive walkthrough method seemed to fit the criteria very well.

As the expert evaluations had provided us some findings about the ease of use of the Internet banking systems, it seemed logical to combine an interview and contextual inquiry to further look into our findings and to understand how the users create and send e-invoices and if they had formed any use habits. The field studies also highlighted problems that the SMEs had come across while using the systems. These two methods were chosen not only to highlight possible usability problems in the services but also to analyze how e-invoicing merged with the traditional way of sending invoices.

2 Electronic Invoicing

In Europe, e-invoicing has a good starting point as the European Union (EU) has been developing the Single European Payment Area (SEPA). SEPA is said to make cross-national electronic payments as easy as current domestic payments [5]. It can also help provide value-added services, such as e-invoicing [6]. The European Council wants e-invoicing to be the primary method for invoicing in Europe by 2020 [7].

2.1 Finvoice Forwarding Service

In Finland, e-invoices are forwarded with the Finvoice forwarding service, which is presented in Figure 1. The four-corner model consists of the sender, the sender's service provider, the recipient and the recipient's service provider. The invoice is forwarded from the sender to the recipient via their service providers. Any material not forwarded to the recipient is returned to the sender. Only a bank or a credit institution can act as a service provider in the forwarding service.

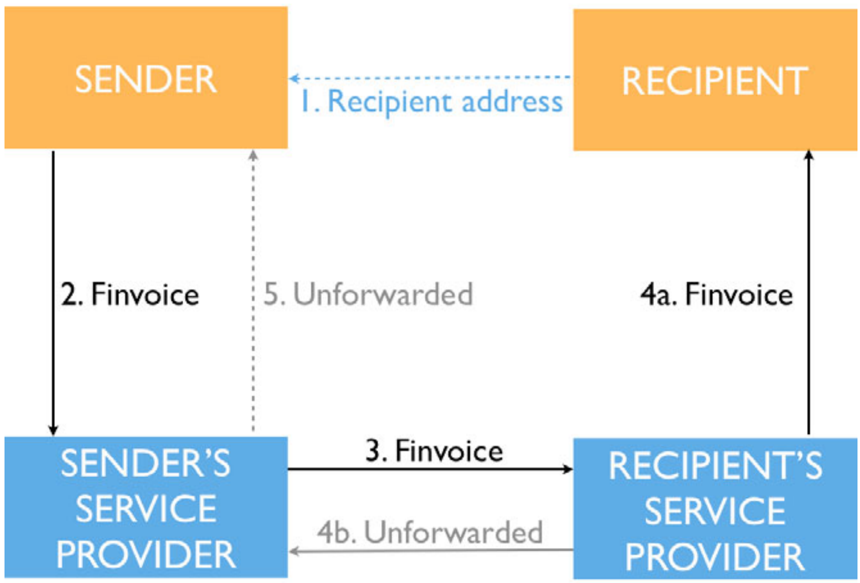


Fig. 1. Finvoice forwarding service [8]

2.2 SMEs and E-Invoicing

For SMEs e-invoicing is problematic as they do not have the skills and infrastructure to implement e-invoicing [9, 10]. Also, the SMEs' attitude can have a negative effect on the adoption of e-invoicing [10]. This is supported by the fact that SMEs' have only a short-term view on IT acceptance and implementation [11].

According to Brun and Lanng, SMEs rely on invoice portals and scanning agencies, which means that these companies do not get the full benefit from e-invoicing. The authors also state that the borders of digitalization should be lowered to provide the full savings benefits for SMEs. The gains for SMEs consist of time and cost savings from the process. The more they send and receive e-invoices, the more benefits the SMEs get. [9]

3 Case: Creating E-Invoices

In this chapter, we present the users, the e-invoice creation workflow and the results of the study.

3.1 Users

A total of 12 users participated in our study. All participants were middle-aged (30-50 years old) people who worked for an SME. They were all in charge of invoicing in their companies. The participants were distributed evenly between genders. Half had used the system for less than six months.

The SMEs were from various industries, such as marketing and communications, web design, floral services, chemical production, knitting machine distribution, graphic design, measurement devices, freelance writing, language services, marine maintenance services and archeology. The monthly number of e-invoices sent varied from one to fifteen. Furthermore, in every SME the number of e-invoices sent was less than half of all sent invoices.

3.2 The Process of Creating an E-Invoice

As the systems follow the Finvoice format, the user interfaces will not be presented individually, but a general description of what information can be included in an e-invoice will be given. The creation of an e-invoice consists of five phases: seller information, payer information, products sold, invoice information and additional information. Moreover, there is always a summary page to scan through the e-invoice before sending it. There is also the possibility of looking at the PDF version of the e-invoice that conforms to the Finvoice format. The process is presented in Figure 2.

Seller Information. The seller information contains a business ID, company name, seller organization unit number, e-invoice address, street address, postal number, place, phone number, bank account number, VAT identifier, contact person, contact person phone number, contact e-mail address and web page. Of this information, the business ID, company name, e-invoice address and contact information are mandatory in the e-invoice.

Payer Information. The payer information contains the e-invoice address, e-invoice operator, VAT identifier, payer name, payer organization unit number, street address, postal number, place, phone number, contact person, contact person phone number and contact e-mail. Of this information, the e-invoice address, e-invoice operator, and contact information are mandatory in the e-invoice. Furthermore, the payers' contact information can be saved in a repository for later use.

Sold Products. The sold products information contains a product name, product code, quantity, unit, unit price, date or time period of delivery and VAT percent. Of this information, only the product code is optional in the e-invoice. Furthermore, the product information could be saved in a repository for use in other e-invoices.

Invoice Information. The invoice information contains the invoice number, invoice date, due date, terms of payment, reference number and order reference number. Of this information, the terms of payment and order reference number are optional in the e-invoice.

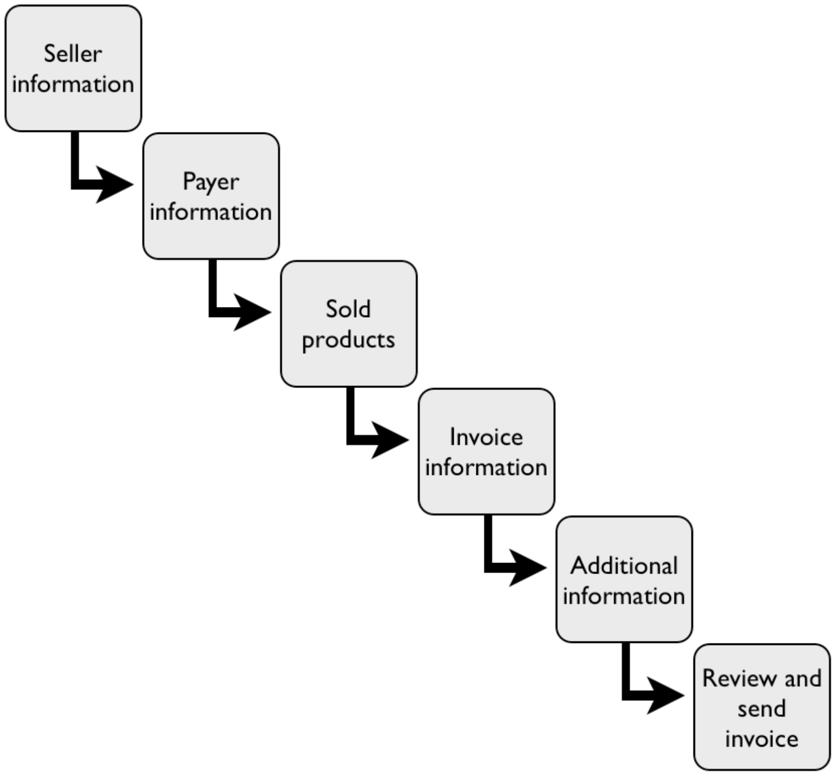


Fig. 2. E-invoice creation workflow

Additional Information. The additional information contains the cash discount percent, cash discount sum, complaint time, penalty interest, additional message, breakdown of costs (URL and text; only one bank had this) and posting information (one bank did not have this). None of the additional information is mandatory in the e-invoice.

Review and Send E-Invoice. The final step for creating an e-invoice consisted of reviewing the invoice and possibly printing out the PDF version of the e-invoice before sending it.

3.3 Results

The following results originate from the expert evaluations and field studies.

Usage of E-Invoices. Eleven of the 12 SMEs had made the transition to e-invoicing because of customer demand. The customer requiring e-invoicing was usually a large company, a Finnish municipality or the Finnish state. In one exception, the entrepreneur running the company knew that he would need to start sending e-invoices soon and therefore had moved his outgoing invoices to the e-invoicing system early on to make the transition smooth.

When asked about the pricing of e-invoices, the SMEs stated that the prices are at an acceptable level. Sending an e-invoice costs €0.25 at two of the banks and €0.85 at one bank. There is also a monthly fixed fee for the service, which for two banks is €5 and for one bank €25.

Although the e-invoicing process needs no paper, often the invoices were still printed out for accounting, which was usually handled in the paper format. One of the interviewees said that e-invoicing does not save paper since he still prints out the invoices for accounting. Nevertheless, for e-invoices his paper usage was half of what paper invoices needed; e-invoices need paper only for accounting purposes.

The Finvoice format has a feature called posting, which means recording a transaction or adjustment to the correct accounts. Using posting helps the e-invoice receiver in its accounting as the transaction is automatically recorded to the correct account. Moreover, posting can be seen as a sign of integrating processes. However, the possibility of adding information for posting was not used by any of the SMEs in the study.

The possibility of saving payer and product information in the e-invoicing system was used in every SME except two: one had only one customer receiving e-invoices, and the other had no fixed product names as the company manufactured only customized products. One of the e-invoicing systems did not have payer and product repositories, but the SMEs used previously sent e-invoices as templates, where they had to change only the sold products, invoice information and additional information fields.

As for sending invoices, the SMEs always sent their invoices after the project had ended. The largest SME with roughly 30 employees sent invoices only twice a month. The others were small enterprises so they could not afford to wait long for the money, but needed it as quickly as possible.

The Finvoice format does not support sending attachments; they have to be delivered by other means. However, the SMEs stated that they need to send attachments.

In the end, the operations method defines if attachments are needed or not. For example, a company selling tangible products can send the needed attachments with the shipped products, whereas a maintenance services company sends the maintenance report after their visit to the customer.

According to our interviews, the SMEs that participated in our study do not need mobile e-invoicing. Usually sending an e-invoice was not so time critical that it could always be done from the office. One company made an exception as their business was running archeological excavations. They did not have an office at the excavation site but sent e-invoices with a smart phone.

Pros. Most important, sending e-invoices is easy. In a normal case, the SMEs can just click through most phases, add a few products and billing information to the invoice and then send it. From a financial point of view, there is no need to buy stamps and letters, which saves money. In addition, there is less need to take letters to the post office anymore, which saves the SMEs some time. For example, one entrepreneur saved a considerable amount of time, as he did not have to drive to the nearest post office, which was 30 kilometers away from his excavation.

The e-invoice was considered quick and reliable. Usually, the e-invoice reaches the payer the next day. According to the SMEs, the e-invoice never gets lost in Finvoice the forwarding service as letters sometimes disappear in the mail system, and as e-invoices are stored in an electronic system, the possibility of a human error in the payer's handling process was less likely. In addition, re-sending the invoice after a problematic situation was a quick process, as the re-sent e-invoice would reach the payer the next day. From the accounting perspective, combining e-invoicing and paper invoicing is effortless as the e-invoice is simply printed out and added to the accounting data.

Cons. In the interviews, the SMEs pointed out also some downsides to the e-invoicing system. The first e-invoice was often sent without any help from the bank, which forced the users outside their comfort zone. They had to learn the process of creating an e-invoice and the system terminology by themselves. The adoption phase was often described as troublesome. For example, the first e-invoice of one SME had a tiny but crucial error, and the company had to cancel the first e-invoice and then send a new one. Some companies also had trouble with payer e-invoice address formats. In the end, the SMEs stated that they were comfortable with creating e-invoices, but the adoption phase and learning to use the system are just something that one has to go through.

The SMEs also stated that understanding and correcting errors is difficult. For example, one of the biggest e-invoicing operators in Finland removes the frame of the e-invoice, which means that some of the routing information is removed from the message.

In this case, the payer e-invoice address was removed, and if additional routing information was not provided, the e-invoice got lost in the system.

As only a minor share of invoices is e-invoices, the SMEs had to also use a paper-based invoicing system. Only one sent paper invoices through banks' mailing services although it was available to everyone. The separate systems introduced problems as the SMEs had to manage two invoicing processes. The SMEs stated that managing two separate systems is burdensome.

The e-invoice is not flexible enough to meet every organization's needs for creating an e-invoice. For example, giving additional information for one product is not possible on the e-invoice. Furthermore, adding a breakdown of costs is not possible in an e-invoice. According to the interviews, the e-invoice is much worse in flexibility when compared to a paper invoice created with a text editor.

In the PDF version of an e-invoice the seller and payer information (A) were in the top-left part, the invoice and additional information (B) in the top-right part, the products (C) in the middle part of the e-invoice and the sum (D) in the lowest. The users stated in the interviews that they did not know where some information would be located in the PDF version, as the process of creating an e-invoice did not match the PDF version's structure. This lack of comprehension caused users not to include some of their information in the outgoing e-invoices. The PDF version of Finvoice is presented in Figure 3.

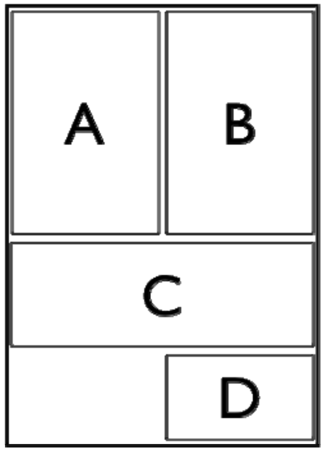


Fig. 3. The PDF version of Finvoice

The fixed costs of e-invoicing services increase the price per e-invoice. Therefore, the SMEs are not able to fully utilize the lower costs of e-invoices, as only a minority of sent invoices are, in fact, e-invoices.

Regarding the terminology, the payer organization unit number caused confusion. The SMEs stated that they do not know what it means or why it is used. Often, it was needed along with the e-invoice address to ensure that the e-invoice reached its destination.

The process of creating an e-invoice was considered to contain needless steps and information. The seller information is the same aside from address or contact person changes. The SMEs pointed out that checking the seller information for every outgoing e-invoice feels unnecessary. Additionally, posting and cash discounts seemed to be unnecessary information that could be hidden from the user interface by choice.

4 Discussion

The emphasis in the research was on the usability and ease of use of e-invoicing. According to the users, creating e-invoices is easy. However, the majority of the users were not technically oriented people. Thus, the low IT skills of the users brought challenges to the process of creating an e-invoice.

All the users who participated this study live in the capitol region. Thus, the sample presents only a tiny fraction of the plethora of Finnish SMEs regionally and lines of business. In our study, only two SMEs conducted their daily work outside their offices. We were not able to interview mobile workers as they were only seldom at their offices. Interviewing mobile workers could have provided valuable information about the benefits of mobile e-invoicing technologies.

Despite our attempts, we were able to interview only people who had already used the e-invoicing system. Interviewing and observing people using the e-invoicing system for the first time could have revealed a more realistic picture of the initial experiences.

5 Conclusions

Against our initial contemplations in the beginning, the study with 12 SMEs suggests that the low adoption of e-invoices in Finnish SMEs and the low utilization rate of opened invoice services are not consequences of low ease of use, a troublesome adoption phase or a mismatch between e-invoicing and established ways of creating invoices. However, the removal of the e-invoice frame by one company lowers the usability of e-invoicing. Overall, the differences between paper and e-invoicing do not explain why the usage is low. Nonetheless, the difference causes extra work for users.

For SMEs, the advantages of e-invoicing are not high enough that they would adopt e-invoicing voluntarily. The more large corporations a company has as its clients, the more it is likely to benefit from e-invoicing. Large corporations get great benefits from e-invoicing, and they are pushing it hard. Consequently, SMEs cannot discard these demands. SMEs cannot get the benefits of e-invoicing due to the low number of companies using e-invoicing. In addition, SMEs have small clients that the companies cannot push into e-invoicing. Although the EU has been promoting the special needs of SMEs, in the current situation they are slipping through the net. To drive the diffusion of e-invoicing in SMEs, the service ecosystem should provide more flexible tools that SMEs can utilize for creating e-invoices suited to their needs.

As our empirical study has provided initial results of the problems that SME face with e-invoicing, a continuation study with a larger sample will be carried out to obtain more concrete results regarding the situation of e-invoicing in Finnish SMEs.

References

1. European Commission Expert Group on e-Invoicing. Mid-Term Report (2009), http://ec.europa.eu/internal_market/payments/docs/einvoicing/report-2009_01_27_en.pdf (referenced January 24, 2011)
2. European Commission Expert Group on e-Invoicing. Summary of Responses (2010), http://ec.europa.eu/internal_market/payments/docs/einvoicing/summary_report_en.pdf (referenced January 24, 2011)
3. Federation of Finnish Financial Services. Finvoice product information, http://www.fkl.fi/verkkolasku/english/electronic_invoice_for_companies/finvoice_product_information.htm (referenced December 22, 2011)
4. Wharton, et al.: The Cognitive Walkthrough Method: A Practitioner's Guide. Technical Report, http://ics.colorado.edu/techpubs/pdf/93-07.pdf
5. European commission. Web page, http://ec.europa.eu/internal_market/payments/sepa/index_en.htm (referenced January 26, 2011)

6. European commission. The single European payment Area – What It Means for Small and Medium Sized Businesses, `http://ec.europa.eu/internal_market/payments/docs/sepa/sme/sme_en.pdf` (referenced January 26, 2011)
7. European commission. Reaping the benefits of electronic invoicing for Europe (2010), `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0712:FIN:EN:HTML` (referenced January 26, 2011)
8. Federation of Finnish Financial Services. Finvoice forwarding service (2010), `http://www.fkl.fi/verkkolasku/english/ladattavat/forwarding_description.pdf` (referenced January 26, 2011)
9. Brun, M.H., Lanng, C.: Reducing barriers for e-business in SME's through an open service oriented infrastructure. In: Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet (ICEC 2006), pp. 403–410. ACM, New York (2006), `http://doi.acm.org/10.1145/1151454.1151519`, doi:10.1145/1151454.1151519
10. Penttinen, E., Hallikainen, P., Salomaki, T.: Impacts of the Implementation of Electronic Invoicing on Buyer-Seller Relationships. In: 42nd Hawaii International Conference on System Sciences, HICSS 2009, January 5–8, pp. 1–10 (2009), doi:10.1109/HICSS.2009.248
11. Sandberg, K.W., Wahlberg, O., Pan, Y.: Acceptance of E-Invoicing in SMEs. In: Proceedings of the 8th International Conference on Engineering Psychology and Cognitive Ergonomics: Held as Part of HCI International 2009 (EPCE 2009), pp. 289–296. Springer-Verlag, Heidelberg (2009), `http://dx.doi.org/10.1007/978-3-642-02728-4_31`, doi:10.1007/978-3-642-02728-4_31

Make Space for the Customer: The Shift towards Customer Centricity

Kostanija Petrovic[1] and Melanie Siegmann[2]

[1] Nokia gate5 GmbH, Invalidenstr. 117, 10115 Berlin, Germany
[2] SAP AG, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany
kostanija.petrovic@nokia.com, melanie.siegmann@sap.com

Abstract. Companies need to understand what their products and services do for their customers. But how can we encourage the organizations that we work for to concentrate more on the needs of their customers and end users? How can we inspire each other to deliver more innovative products?

In this paper, we will discuss the impact of Design Thinking, facilitated by User Experience Design (UxD) on different levels and divisions of an organization and show how UxD can help support an organization's shift towards customer centricity.

Keywords: Innovation, Design Thinking, Customer Experience, User Experience.

1 Introduction

Peter Drucker famously stated that "the single most important thing to remember about any enterprise is that there are no results inside its walls. The result of a business is a satisfied customer". [1] This assertion is just as valid today as it was back then. And user experience in enterprise software is probably the field with most potential for improvement regarding customer experience and satisfaction.

Within the last decade, many large software companies such as Microsoft and SAP have introduced common user-centered design processes in their organizations, aligning these methods in their respective software engineering processes. This development shows that User Experience is indisputably a critical quality factor in software development, and it acknowledges that customers expect effortless, engaging interaction with desktop and web applications.

As a user experience designer, you have probably sat through numerous meetings where someone involved in your project expected you to make things "nicer looking" or "easier to use". Maybe someone was discussing innovation in the context of "really cool" features, like the Cover Flow technology in Apple iTunes® for example. Most likely you have sighed and felt that, once again, your profession is not being taken seriously. As professionals we know that creating software that is innovative and generates an enjoyable user experience is more difficult than many companies and stakeholders appreciate.

So what is innovation and how does it relate to design thinking? According to Schumpeter [2], innovation is the commercial or industrial application of something

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 485–490, 2011.
© Springer-Verlag Berlin Heidelberg 2011

new, a new product, process or method of production. He makes a distinction between innovation and invention, as the latter merely encompasses new ideas, whereas innovation involves the commercial application as well as the new idea.

Design Thinking [3] is a methodology that facilitates the identification of user and customer needs and the creation of potential solutions addressing these needs. It can also be seen as a creative problem-solving approach that can be applied to all kinds of problems. According to Brown, the cyclic, iterative process that is always a team activity encompasses the three phases Inspiration, Ideation, and Implementation. During the Inspiration phase, the team defines the problem and conducts research by observing and interviewing members of the target audience. The insights are shared in a group setting, ideally a project room. In the Ideation phase, the team brainstorms all possible types of ideas and creates prototypes that are validated and iterated. Once the team has agreed on the solution approach, the Implementation phase starts; this is when the solution is turned into a real product or service.

The Design Thinking approach [3] has been around since 2008, and looking closely, it doesn't seem that difficult to apply. While this approach works very well at IDEO, the company that Brown invented the concept for, working in a large enterprise setting has its own special challenges.

2 Impact on Different Organizational Levels

How does design thinking work in a large enterprise? Several factors must be considered when promoting design thinking within the enterprise to ensure the delivery of a product that fits the needs of customers and end users and, according to Schumpeter [2] provides value to customers.

Let's start at the executive level. As we all know, it is crucial to campaign for executive support because board members have to understand the importance of User Experience Design as a strategic engagement. We can achieve this by demonstrating the value of UxD within selected strategic projects that receive respective funding and resources. We also have to follow certain best practices, such as validating designs early and often, and making the results accessible throughout the organization. Most executives are indifferent towards the particulars of our profession, but are interested in how we can help create revenue. So we have to rephrase our message and emphasize the benefits; usability issues are of secondary importance here.

It is still surprising how many UxD teams have limited or no exposure at all to their sales and marketing departments. However it is precisely these people with whom we need to share our results; it is important to emphasize the value added for customers and end users and the direct relationship between this added value and an organization's strategic objectives [4, 5].

All of the above can help to increase visibility. However, the executive level is in charge of promoting a corporate design culture and together with upper management levels, they need to provide the guidance that encourages employees to embody this vision and act upon it. This needs to be a mandate and requires cultural change.

On an operational level, UxD quality criteria need to be an integral part of the product development process.

In many cases, user experience designers undertake quality assurance work in development projects and find that, once they identify usability issues, these are not acted upon; problems that they discover may not be easily solvable due to technical restrictions or other important features that would otherwise be compromised. As a consequence, the product remains hard to use. To counter this, we need to establish that:

- UxD is part of the product definition phase [6]
- Product concept including information architecture and product design vision is defined before development starts
- Nothing is implemented that does not meet known and proven customer needs
- Following UxD specifications is a major requirement for development colleagues
- The assignment of UxD staff to projects follows project management best practices (such as assigning pairs of usability experts to one project instead of splitting individuals over multiple projects)
- Usability is part of the quality assurance procedure and major usability defects have to be fixed prior to release

Irrespective of the development method (either traditional or agile) User Experience resources need to be involved in projects from the very beginning. If this is not possible, they'll only be there to limit the damage. This pre-phase can be a funding gate that cannot be crossed without executive sign-off before real coding starts [6].

Moving on to the next organizational level down, we'll now look at the UxD managers, the ones that assign UxD resources to various development projects and take the decisions that affect the design staff on a day to day basis. The biggest challenge for UxD managers is to ensure that user experience designers are embedded in projects where user experience is likely to have the greatest impact on product success and thus provide the most customer value.

Nieters et al. [7] offer a good set of criteria that help to identify such strategic projects:

- Product team receptivity. The product team itself requests support from UxD, rather than have it imposed upon it by management. If a product team is ambivalent, the UxD group disengages
- Potential revenue or cost savings. The UxD group seeks projects that they anticipate will see a significant revenue increase in the first year
- Advanced or emerging technology – a new technology that has not yet been introduced to the market, so the UxD group can make a larger impact than on legacy products
- Leveraging the respective user experience standards (user interface guidelines). If a product team does not intend to adopt these standards, the UxD group will not assign resources
- High visibility. If a project is a "pet project" of a cross-functional or highly visible organization within the company, the UxD group is more willing to accept it

- Point in the product lifecycle. If design has already begun, it is often too late to impact a product's overall experience on a fundamental level. There are times when the UxD group agrees to work on a project through multiple iterations, starting late in one cycle to impact a subsequent release
- Realistic time-to-market demands. If project schedules make delivering a high-quality user experience impossible, the UxD group is less likely to accept the project.

The UxD managers should also act as facilitators for cultural and process change. They should be assessed by the value of customer insights that get used, adopted, and implemented across the organization [8].

However, it's not just the organization around us that has to change; we as user experience professionals also need to change in order to help transform the organization in way that we call "making space for the customer".

Every member of a company's User Experience Design team needs to shoulder responsibility for satisfying its customers. User experience professionals can glean customer insight by employing user-centered design methods and carrying the results through the product development process. This helps to promote customer centricity by facilitating the adoption of user-centered design methods, tools, and deliverables and encouraging fellow team members to do likewise.

In this regard, SAP has introduced the role of Ux Advocate in product and Scrum teams. Colleagues from development and product management are taking over Ux tasks, helping to scale the user-centered design process within the context of agile development. The Ux Advocates are embedded within a team of ten and receive coaching from members of the Ux team, ensuring that they ask the right questions: who are the end users, what goals are they working towards, and what tasks do they perform?

In addition, the recently established customer engagement initiative (CEI) – enables SAP project teams to engage more easily with customers in all phases of product development, from invent to implement – helps drive the shift towards customer-centricity.

It's not just actual product designs created in the design process that benefit the customer experience. Many other artifacts are generated that can help market a product successfully. Establishing a persona and scenario library, has, for instance proven a useful practice at OpenText; all relevant stakeholders, such as product management, sales enablement, marketing – and even training services – can search for and find information. OpenText uses its personas consistently from development through to marketing of products, including the definition of demo scenarios for new software releases. This is consistent with Watson [8] from Wells Fargo stressing that it is crucial to keep real people and their stories at the center of product development. Wells Fargo accomplished this by creating consumable and reusable UCD tools – key tools such as user profiles and user task models, actionable by their connection to so-called "scenario starters" – to communicate with different audiences. Instead of formal training sessions, Wells Fargo's Ux team shared methods across the organization, allowing stakeholders to engage with these reusable UCD tools without recourse to the design team.

3 Conclusion

User Experience Design is a strategic topic that interfaces with many other divisions in an organization. Generally perceived as a service provider to development, it is actually a source of insights about customer needs, insights that are the basis for creating great products, customer, and end user experiences. Deploying UxD across the enterprise requires management support, strategic engagements, and cross-departmental collaboration. It also requires us as UxD professionals to think and act strategically. Customers should continue to demand better experiences that will help us to change the company culture and structure.

When we talk about "making space for the customer" we are referring to space in the heads and hearts of our colleagues. After all, it takes an entire company to deliver a great user experience.

Based on our own experience, and borrowing from the relevant published recommendations [3, 9], we have identified the following set of principles that you can implement to make space for your customers:

Make yourselves easy to use! Make it easy to consume as well as to refer to customer insights. Make your deliverables stakeholder-centered. Show how the insight fits in the process, create empathy by letting people feel what frustrates as well as what delights customers and end users; show them, instead of telling them.
Get there early! When working on new projects, aim to get there ahead of time – be involved from the beginning, during the research and ideation phases. Be a facilitator, and provide access to customer insights.
Create a vision! Describe how becoming customer-centered encompasses multiple product aspects. Look at what is needed in a project and learn from your stakeholders, and let them learn from you, but without causing bad feelings or setting up communication barriers.
Leave your silo! Cross collaborate with Sales, Marketing, Product Management, as well as development and be willing to share and gain insights. Work with the customer intelligence team or the colleagues at your company who measure customer satisfaction.
Get strategic! Refer to User Experience Design as a strategic engagement that should be reflected on different organizational levels, and look and campaign for executive support.
Give and share credit! When a project is successful, make sure to be there to take and share the credit, as the value provided by user-centered design needs to be made transparent to the audience. They like your product – excellent, let them know what made it so great. Share the credit, again, across the organization.
Measure success! User experience professionals should get assessed in their capacity as facilitators, the value of insights that get used, and methods that are adopted across the organization to drive cultural and process change should be appreciated and communicated.

References

1. Drucker, P.F.: Peter Drucker on the Profession of Management. Harvard Business School Press, Boston (2003)
2. Schumpeter, J.: The Theory of Economic Development. Harvard University Press, Cambridge (1934)
3. Brown, T.: Design Thinking. Harvard Business Review 13(6), 84–92 (2008)
4. Lund, A.M.: Usability ROI as Strategic Tool. UPA User Experience Magazine 6(2), 8–10 (2007)
5. Siegel, D.A.: The business case for user-centered design: increasing your power of persuasion. ACM Interactions 10(3), 30–36 (2003)
6. Rosenberg, D.: Garbage In, Garbage Out – the Agile Way. UPA User Experience Magazine 9(1), 32 (2010)
7. Nieters, J.E., Ivaturi, S., Dworman, G.: The Internal Consultancy Model for Strategic UxD Relevance, CHI 2007, San Jose, California, USA, pp. 1825–1832 (2007)
8. Watson, S.: The Business of Customer Experience: Lessons Learned at Wells Fargo. ACM Interactions 15(1), 38–43 (2008)
9. Heath, C., Heath, D.: Switch: How to change things when change is hard. Broadway Books, New York (2010)

The New Experience for Business: Why User Experience Is the Differentiation Strategy in the Cloud Context

Andrey Sirotkin and Bronan McCabe

VTT Technical Research Centre,
Kaitoväylä 1, P.O. Box 1100, Oulu, FI-90571, Finland
{andrey.sirotkin,bronan.mccabe}@vtt.fi

Abstract. Could technologies transform software industry and drive the salience of user experience as differentiation strategy. However the current product augmentation paradigm conditions company's ability to design and market an experience product. Without new perspective on product augmentation a company may not succeed in communicating the value of user experience to its customers. Thus the differentiation strategy may be jeopardized. The paper proposes the transformation of traditional view of augmentation. The new model reflects the need for designing user experience prior to product or service development. Companies must build new capabilities to build a successful user experience differentiation strategy.

Keywords: user experience, differentiation, product augmentation, cloud computing.

1 Introduction

Despite many challenges for advancement of Cloud technologies [3, 5], Cloud computing shapes the way software is developed and marketed [7, 16]. The developers do not need large capital resources to implement innovative ideas. Scalability of cloud services eliminates the risk of in-house over- or under-capacity of computational power, thus minimizing the threat of overinvestment in hardware or lost of potential revenue. Furthermore, the cost structure changes as software becomes a service rather than an application sold for an isolated environment. Pricing models become a key element for reaching vast markets with economies of scale allowing to set zero or close-to-zero prices [1]. As service-based software undergoes standardization to support reuse of elements in different environments and scalability of services [2] it furthers the commoditization of software products.

As a result, the barriers of competition in software industry change. Innovative ideas can find their way to market utilizing Cloud resource. It is a growing challenge to be differentiated from growing competition. In Cloud context value of product or service as perceived by customer becomes a key differentiating factor. This draws interest to user experience as a differentiation strategy because it can provide an opportunity to add value to customers. That said, communicating the added value is a challenge primarily because, if it is not distinguished from aesthetics and usability then customers are unlikely to perceive it as differentiating and of consequence, are unlikely to be attracted or pay anymore for it.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 491–499, 2011.
© Springer-Verlag Berlin Heidelberg 2011

A software company which seeks to differentiate its products or services must recognize that it cannot achieve the desired results by simply adding more services to an existing product. A differentiating experience is created in the way a product or service is conceived [6] and marketed. For many software companies this means that should they wish to transition into the experience market [11] they will need to consider new capabilities for supporting it. Whilst aesthetics and usability design is important, on its own it is likely to be insufficient in the new business arena of experience. Failing to recognize the gap between current design processes and those required for developing experience will likely result in doing-more-of-the-same, and failure to differentiate.

2 Experience Enabled by Cloud

By definition experience is personal and unique, meaning the ability to stage a unique and memorable event is fundamental for experience creation [11]. This is flexibility of a different kind and it hinges on competence development, i.e. providing enough flexibility in organizational processes to enable consumption by an audience of one. However, organizations are driven by the production of goods and services that are created for mass consumption. This tension is creating a paradigm shift; one in which organizations have to balance the need for mass production with the consumers desire for personalization [13].

A distinct characteristic of the user experience business is the opportunity for the users to create their own experience. The product which a customer gets is the experience that he or she creates in partnership with the company. It is different from customization or user-centered design, which employs methods for producing a final product targeted at satisfying known consumer needs. An experience product is a result of value co-creation and is targeted at satisfying customer values.

Human values are known to be universal [14] but they are understood, expressed and attended to differently depending on context and situation [17]. Thus understanding customer values, contextualizing them and designing an experience targeting values all become a pivotal principal of success.

Creating and delivering experience offerings requires advanced understanding of criteria individuals use to evaluate experiences against their standards and preferences. Cloud technology makes such comprehension possible. This means that companies that operate in cloud business are well positioned to take advantage of the experience business. Cloud business benefits from access to a wide range of consumers, on the one hand; and an ability to accrue a deep sense of customer desire at an individual level on the other. Thus Cloud can be one of the means for a company to utilize its resource base to deliver a unique experience for every customer as suggested by Prahalad and Krishnan [13]. In other words, cloud technology makes tailoring experiences for one customer possible.

Although tourism, entertainment and other service businesses have been long exploring the idea of an experience offering, the cloud phenomenon opens up the experience market for the producers of all kinds of goods and services. Cloud companies can appeal to customer values on a personal level, with great accuracy. Corporate management can leverage new internet-based technologies to align internal system architecture with the life events of individual customers and thereby opening up new prosperous business opportunities.

3 Research Focus and Methods

Experience design is subject to much debate in business and academic communities (e.g. [8]). This paper used case-base research approach to understand the issues underlying a company's weakness in differentiating its experience products from other software products. The goal of the study was to suggest ways to enhance the competitiveness of the user experience business in the context of Cloud. Case-based research can be very effective when reviewing challenges in the existing environment [18].

The paper is based on a study of a Finnish business unit in a multinational company. It included interviews of middle management in the unit. Soft system methodology [4] was utilized in structuring and analyzing the interviews. Additionally freely available information from web-sites and publications was used for culture web analysis [9] and core competence model [12]. The findings initiated the review of the current product augmentation paradigm [10]. Since case-based research can be used to initiate the development of a theory [8], this paper shows how the current product augmentation view must be modified in order to reflect the process of designing an experience offering.

3.1 Data Collection and Analysis Methods

The interviews consisted of two parts and a total of five managers were interviewed. During the first part each interviewee was asked to draw elements relationships and conflicts of user experience rich picture. Later all pictures were combined into a common view. This approach was selected for two reasons. First, it gives an opportunity to build a holistic perception of user experience while avoiding extensive debate over specific terms. Second, it unites employees with different backgrounds and combines their diverse views around one shared view. The other half was a semi-structured interview regarding the company's perceived strengths and weaknesses.

Soft system methodology [4] was used to establish the current and desired views of user experience. Root definitions for both views were developed with the help of CATWOE method [4]. The comparison of the currently perceived and future company positions identified gaps and generated suggestions about possible changes. Prahalad and Krishnan [13] suggest that such analysis is necessary for building new corporate capabilities. The current business situation was described by the following definition.

A system owned by *the company* and carried out by its *experts* via the *software development* process – produces *unsatisfied customers* because the company's worldview is *technology expertise* and is constrained by the *understanding of end users*.

The same method was used to develop a definition of the desired future.

A system owned by the *company* and targeted at *user experiences* creates happy customers because it utilizes understanding of end *customer values* where *experience* is the worldview.

By comparing the elements of the two definitions the company can plan steps for creating capabilities for desired future [13]. For example, the company must alter its transformation process in order to produce satisfied customers. However the worldview conditions the changes that could be made in the transformation process. The new world view has to be grounded in understanding what customers value in software experience. This understanding is expressed in the way management thinks about the experience business. Thus transforming the way the company thinks about its business would facilitate the desired change.

Culture web analysis and core competence model were used to refine the understanding of the perceived situation and the depth of required modifications. The product augmentation model [10] was used to summarize the change that the company has to make in order to build capabilities for experience market.

The overarching theme of the findings was that the company is in the business of developing software rather than designing experience. User experience is perceived as designing the "look and feel" and is not aimed at customer values. Development of software for corporate partners needs to include more of the end user perspective on the product. This is because corporate partners do not value technology per se but rather view it as means for satisfying end user needs. For the company to develop a user experience competence it must serve its B2B partners with awareness of end user values [13]. That is, the company has to understand what kind of experience it wants to provide to end users.

Findings of the research were presented and discussed with the management. As a result some immediate changes were made to the development process with a greater focus on end users. The midterm goals were set for improving the multidisciplinary team competence. Also a more extensive follow-up study of user values in the Cloud context is planed. In the long term the company plans to incorporate the understanding of end user values in the design process. This will help the company to continue its transition from the software developer to an experience designer.

The generalized principles of designing experience are discussed in the following section.

4 Designing Experience

The case study proposed a broad assumption that there are two generally held perspectives on user experience. The first one focuses on customers' goals, needs, behaviors, and emotions when using a product (e.g. [15]). A company surveys various market segments and incorporates the feedback to improve its products. This is a traditional approach that perceives user experience as an incremental product improvement. The challenge with this view is that experience is only perceived at a product level.

As it was discussed above Cloud technologies shift perceive customer value from tangible product towards an intangible experience. The case-study company was being affected by such an external force. A failure to recognize the changes initiated by Could may lead to overlooking the new opportunities and instead being threatened by industry drivers.

The company in our research may be representative of some interactive technology producers who adopt the above view of user experience. Such a company may choose to evaluate user experience by measuring how an offering fulfils customers' needs and wants. This perspective gives the ground for setting new targets and measuring the magnitude of user experience. As a result a product is being gradually improved as more observations of the product usage are collected. However such an approach is questionable. Henry Ford is known for saying, "If I'd asked my customers what they wanted, they would have said faster horses." Hence a company that aims at satisfying only known expectations is unlikely to become a key influencer in a market. Moreover, in waiting for feedback the company risks always being one step behind the industry innovators. Market studies are still important; however, their interpretation should be adjusted to take into account that customers may not know any better ways of satisfying their needs than in the current context. Additionally, they may not suggest methods for altering the context, either.

The companies that put a product in the center of their design process view user experience as an incremental improvement of the existing product. Figure 1 illustrates the conventional business view of product augmentation [10]. At the core, the product appeals to a known need and the actual product is designed to satisfy it. Services are then added to the actual product to provide some level of differentiation. User experience is viewed as adding another augmentation layer for achieving more differentiation. Corporate management treats experience product in much the same way as it does the other layers.

The challenge with this approach is that current business processes may fall short of designing for fun, enjoyment or pleasure. The problem is compounded when trying to add other experience qualities such as hedonics. Though such qualities are not inherent of a product, they are attributed to it during the experience of it. Because actual and augmented layers are both a part of user experience, all blur together in the minds of both designers and customers. Thus, it becomes very difficult to communicate any added value proposition.

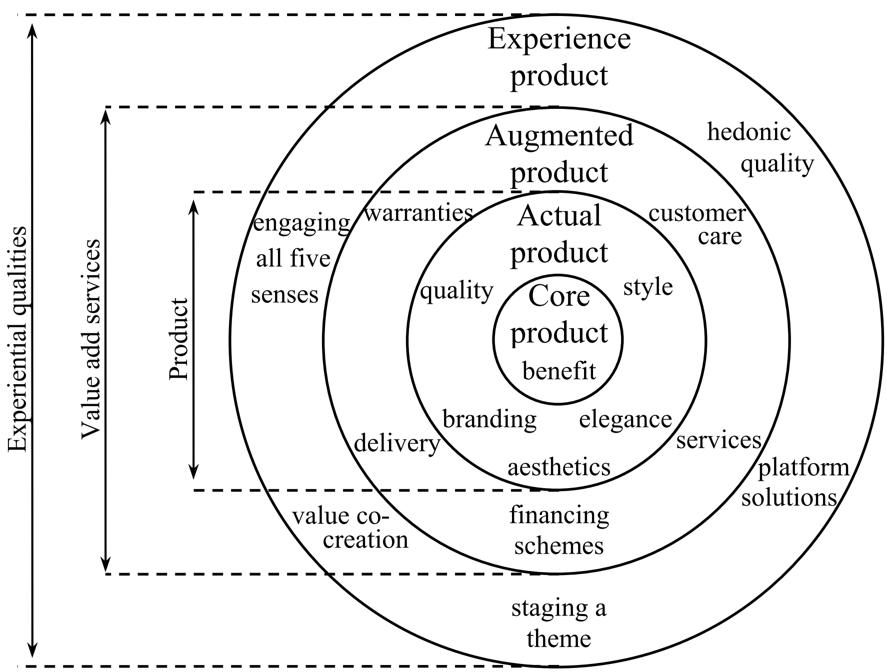


Fig. 1. Conventional business view of product augmentation

There is a noticeable difference between an organization that sells actual products and its competitor who provides augmentation, i.e. customer support and a warranty in its offering. The two companies compete at a different level, enabling one to differentiate itself from the other. However, in order to provide this differentiation and move from an actual to augmented product offering, an organization must develop certain corporate capabilities. This is similar to the transition an organization must make to provide an experience offering.

The second perspective and a different view of user experience is when an organization emphasizes the competencies that enable design for experience in the first place [6]. In other words, corporate management aims at certain kind of experience and designs an offering that supports it. This perspective creates a need for business change, which begins with the transformation of the company's viewpoint and focuses on the development of capabilities. The challenge is to combine diverse resources and deliver personalized experiences in all business areas that a customer comes in contact with.

The experience business model differs from the traditional one. It is grounded in the principle of staging an experience [11]. A companies' production process brings together talents, resources, goods and services in a coherent whole that engages individual customers. Consumption becomes a co-creation of experience, in which both customers and a company are involved [13]. The resulting take-away product is a memorable event. This is a competition imperative. This means that corporate management should view all the company's outputs, processes and resources as means for making a platform where customers can create their personal experiences in partnership with the organization. In designing experiences there is a great need for this new mindset. Re-evaluation of the business viewpoint and re-alignment of the capabilities and strategies are the prerequisites of offering user experience as a product.

The focus of design shifts from an attempt to make a hedonic product to staging an experience. The goal of market research also moves from surveying needs to understanding the value of experiences also. When values are contextualized they can be expressed as needs. Knowledge of values enables a company to shape the context around them and empower the customers to have their diverse needs met as they engage with a product on a personal level.

Contrary to the conventional product development view the step that follows value analysis is not actual product but experience design. Since needs are subjective and context dependent it is sensible to understand how the company will shape the experience. Therefore, any organizational offering should start with answering the question, "what kind of experience should this product or service be designed to deliver". Experience becomes a central concern and the product is designed around it. The value of the product is attributed to it by the customer either during an experience or a series of experiences. It is an attribution process that makes the product highly valuable to the customer.

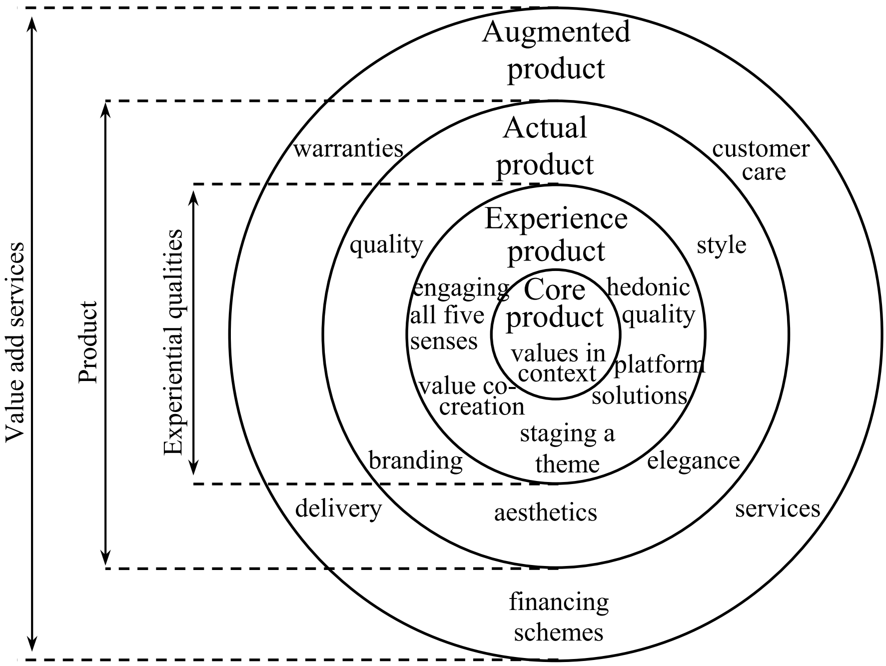


Fig. 2. Paradigm shift for experience product

As a result of the changed business viewpoint an organization makes not product, but experience central to its development process. Figure 2 summarizes the paradigm shift of product augmentation. Experience becomes central to the design offering with actual product and additional services supporting the experience design. The value is cocreated by customers as they use a product in various situations. The support for such a perspective requires an organization to develop a strategy for investing in corporate capabilities. These new capabilities support the transition to the next level of differentiation, which is an essential ingredient for gaining competitive advantage.

5 Conclusion

While the user experience business may seem like an incremental step for software integrators it is important to recognize the extent of change required to embrace the opportunity. Based on the perception of the current situation in the case study an experience that a good software developer provides may not be the one its customers value. Thus it can be questioned whether the company is truly designing experience. It may be that there is some confusion between designing experience and elegance. As a result a company's ability to take advantage of the experience market is at best limited. However, if a clear distinction between experience offering and usability or aesthetics is made (i.e. the desired and current states), a company has the ability to produce sustainable added benefits and communicate them well.

Conventionally organizations focus their efforts on actual product design. At VTT, we suggest that the foremost task is to focus on the experience for which the product is being designed. This can be achieved by studying customer values and contextualizing them for a particular market or context. Cloud technologies can enable companies to accrue knowledge about experiences that their customers value. Understanding the experience allows the company to develop a strategy for building user experience competence. In the Cloud business landscape user experience provides the differentiation that enables companies to compete, tolerate uncertainty and gain competitive advantage.

References

1. Anderson, C.: Free: The Future of a Radical Price. Random House Business Books, London (2007)
2. Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., Brandic, I.: Cloud Computing and Emerging It Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility. Future Generation Computer Systems 25(6), 599–616 (2009)
3. Carr, N.: The bid switch: Rewriting the World, Edison to Google. Norton, New York
4. Checkland, P.: Systems Thinking, Systems Practice. John Wiley & Sons, Chichester (2009)
5. Dikaiakos, M.D., Katsaros, D., Mehra, P., Pallis, G., Vakali, A.: Cloud Computing: Distributed Internet Computing for IT and Scientific Research. IEEE Internet Computing 13(3), 10–13 (2009)
6. Eisenhardt, K.M.: Building Theories from Case Study Research. The Academy of Management Review 14(4), 532–550 (1989)
7. Hassenzahl, M.: Technology for All the Right Reasons. Morgan & Claypool, San Rafael (2010)
8. Hassenzahl, M., Tractinsky, N.: User Experience - a Research Agenda. Behaviour & Information Technology 25(2), 91–97 (2006)
9. Johnson, G., Scholes, K., Whittington, R.: Exploring Corporate Strategy, 8th edn. Person Education, London (2008)
10. Kotler, P.: Marketing Management. 11th international edn. Person Education, Upper Saddle River (2003)
11. Pine II, B.J., Gilmore, J.H.: Welcome to the Experience Economy. Harvard Business Review 76(4), 97–105 (1998)
12. Prahalad, C.K., Hamel, G.: The Core Competence of the Corporation. Harvard Business Review 68(3), 79–91 (1990)
13. Prahalad, C.K., Krishnan, M.S.: The New Age of Innovation: Driving Co-created Value through Global Networks. McGraw-Hill, New York (2008)
14. Schwartz, S.H.: Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. Advances in Experimental Social Psychology 25, 1–65 (1992)

15. Tullis, T., Albert, B.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Morgan Kaufman, Burlington (2008)
16. Vouk, M.A.: Cloud Computing – Issues, Research and Implementations. In: Information Technology Interfaces, ITI 2008, 30th International Conference on Information Technology Interfaces (2008)
17. Williams Jr., R.M.: Change and Stability in Values and Value Systems: a Sociological Perspective. In: Rokeach, M.: Understanding Human Values: Individual and Societal, pp.15-46. The Free Press, New York (1979)
18. Yin, R.K.: Case Study Research: Design and Methods, 4th edn. Sage, Los Angeles (2009)

Managing a Green Giant: Sustainability Performance Management, a User Experience Perspective

Ben Tomsky and Angad Manco

SAP Labs, LLC, USA
{benjamin.tomsky,gurangat.manco}@sap.com

Abstract. With a backdrop of global resource depletion, climate change, and environmental degradation, there's growing pressure on businesses (especially big business) to change course, to conduct themselves sustainably. Given SAP's mission to not just sell software, but to "help the world run better," it is well positioned to advance the cause of sustainable business. This paper summarily defines the topic of Sustainability Performance Management and the objectives of a business software product designed for the space. In greater detail, it addresses the research methods, design process, and development collaboration model employed during the development of SAP Sustainability Performance Management (SuPM).

Keywords: Sustainability, Performance Management, Operational Performance, Key Performance Indicators, Key Performance Drivers, Management Framework, GRI, DJSI, CDP, Analysis, Reporting, Agile.

1 Introduction

"*Sustainability is increasingly mission-critical across the corporate world… It is becoming a key element of business strategy, with the potential to deliver new levels of efficiency and operational performance, innovation and organizational transformation.*"

- Dr. Stephen Stokes, VP of Sustainability & Green Technologies, AMR Research

Perhaps it's increasingly mission-critical… but what exactly is "Sustainability" in the "corporate world"? Perhaps the best definition comes from analogy to existent business practice, by way of characterizing sustainability as "triple bottom line accounting": triple bottom line accounting means expanding the traditional concepts around performance and reporting to account for not just financial performance, but also ecological and social performance.

Returning to Dr. Stokes: "Managing and reporting an organization's sustainable performance via transparent and high quality data collation, analysis, optimization and modeling is a new basis for defining and communicating operational excellence." As is the case with other measures of operational performance, such a new basis of operational excellence will undoubtedly rely heavily on business software. SAP organized around this opportunity in 2008, and established the foundation for what was to become SAP Sustainability Performance Management (SuPM).

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 500–508, 2011.
© Springer-Verlag Berlin Heidelberg 2011

SAP SuPM was envisioned to support the following objectives:

- Incorporate social and environmental issues and data into mainstream business practices and processes
- Reduce the cost and increase the reliability of gathering environmental and social performance metrics
- Reduce the complexity of preparing reports based upon multiple different sustainability reporting requirements
- Make sustainability actionable
- Improve visibility of sustainability-relevant key performance drivers
- Support reporting at relevant levels of the organization (e.g., plant/building, country, region, global)
- Provide a management framework for guiding improvement

In this paper, we will review the research methods, design process, and development collaboration model employed for Sustainability Performance Management.

2 Research Process and Deliverables

User research activities aimed to gain an understanding of the high-level business process, customer pain points, user profiles, as well as key tasks.

The nascency of the Sustainability Performance Management domain presented research challenges, insofar as best-practice business processes have not been established within industries, much less spanning industry categories. While NGOs such as the Global Reporting Initiative (GRI), Carbon Disclosure Project (CDP), and Dow Jones Sustainability Index (DJSI) provide guidelines and frameworks for publicly disclosing aspects of sustainability performance, they fall short of prescribing specific internal processes to collect, manage, and act on sustainability metrics.

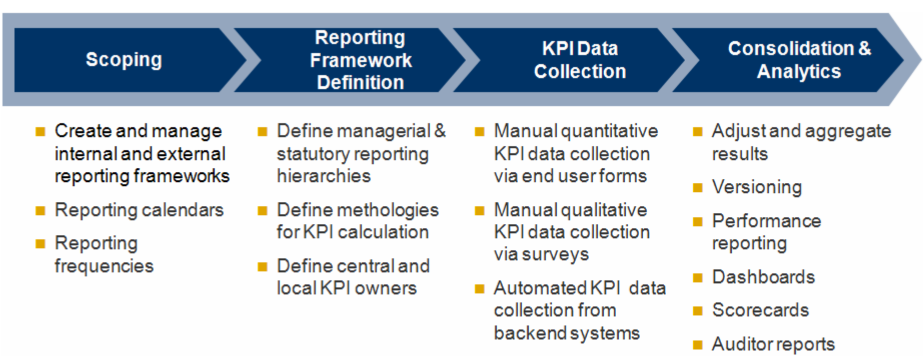


Fig. 1. SAP SuPM process phases associated tasks

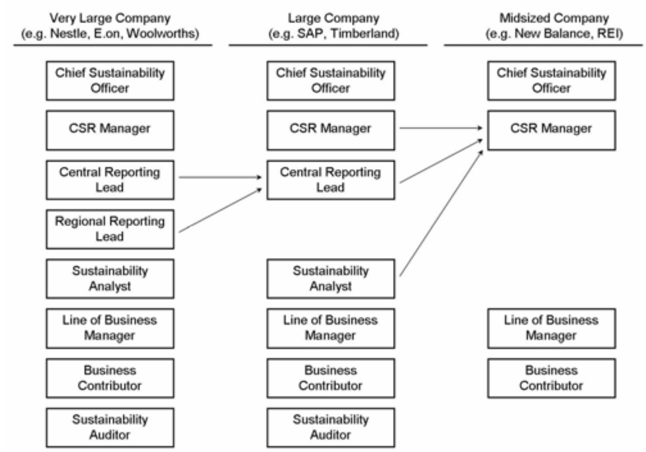


Fig. 2. Role consolidation correlated to company size

Table 1. Excerpt from SAP SuPM 1.0 Role-Task Matrix

	Candidate for Integration or Reuse	Chief Sustainability Officer	Sustainability Team Lead	Sustainability Team Member	Business Contributor
Scoping					
Create/Upate/Change Stakeholder Categories			x		
Select reporting frameworks and application levels		x	x		
Create new reporting framework and assign candidate KPI's		x	x	x	
Create stakeholders			x	xx	x
Create/update stakeholder profile				xx	x
Define process for application levels of new reporting framework			x	x	
Define business contributors for a stakeholder risk assessment			xx	xx	
Issue stakeholder risk assessment		x	x		
Respond to stakeholder risk assessment (and create stakeholder profile)		x			x
Review and adjust stakeholder risk assessment results		x	x		
Create/edit stakeholder survey				x	
Issue stakeholder feedback survey			x		
Respond to stakeholder feedback survey					xx
Review stakeholder feedback and finalize report scope (materiality analysis)		x	x		

Taking an opportunistic, breadth-before-depth approach to the customer and user research, SAP User Experience (UX) defined a research plan consisting of business process walkthroughs, remote phone-based interviews, as well as contextual interviews. In the initial product definition phase, we worked with eight companies to arrive at and refine a standard process for sustainability performance management (with an orientation mostly toward disclosure, at this time), along with the primary features/functions and tasks associated with each phase.

User-level research revealed the key roles involved with the process, as well as how the roles vary in proportion to an organization's size. The key roles indentified across all companies, regardless of size, were the Chief Sustainability Officer, Corporate Social Responsibility (CSR) Manager, and the Line of Business Manager. In larger companies, it was common to find greater specialization and differentiation, with roles such as Sustainability Analysts, Central Sustainability Reporting Lead, and Regional Reporting Leads. In order to define a role-based user experience that would easily scale within the observed variability, we mapped how role-based responsibilities consolidated as company size shrank:

Combining customer-level business process research with user-level task research yielded the role-task matrix, the foundation on which the role-based information architecture could be designed.

3 Design Activities and Collaboration Model

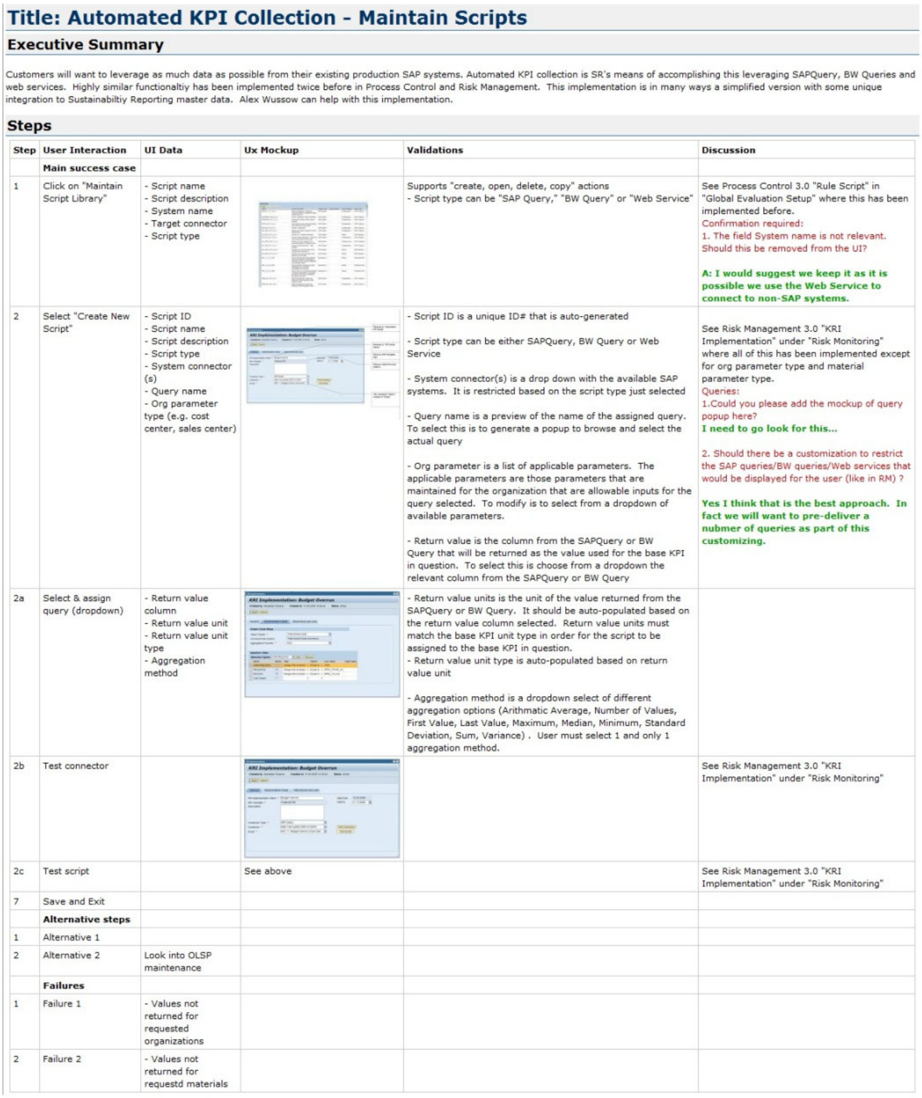
Based upon the market and customer knowledge gleaned during the research phase, the leadership in Product Management determined that a version 1.0 release within the calendar year 2009 was essential to establish SAP's place in the incipient market. Such a release-date driven schedule contributed to the team's decision to employ Agile Development methods for the release, an atypical approach within SAP at the time. We embraced a fixed-timeline, variable-scope mentality, and scheduled eight three-week development sprints.

A globally distributed team was formed, consisting of:

- Two Solution Managers in Palo Alto, California, with an assistant later joining in Walldorf, Germany
- A team of twelve Developers and QA Specialists located in Bangalore, India
- Two UX Designers in Bangalore, with a Principal UX Designer in Palo Alto
- Consulting Development Architects in Walldorf

Detailed product definition and UX design activities were aligned with development sprints such that requirements, use cases, and UX designs were prepared exactly one sprint before development would build the features. Within the inbound product definition work stream, each sprint included:

- Co-design (between Product Management and UX) of interaction flows, wireframes, and high-fidelity mockups—a highly collaborative design process that often consisted of half-day working sessions
- Publication of use cases, interaction flows, and mockups on an internal wiki.



Title: Automated KPI Collection - Maintain Scripts

Executive Summary

Customers will want to leverage as much data as possible from their existing production SAP systems. Automated KPI collection is SR's means of accomplishing this leveraging SAPQuery, BW Queries and web services. Highly similar functionaltiy has been implemented twice before in Process Control and Risk Management. This implementation is in many ways a simplified version with some unique integration to Sustainabiltiy Reporting master data. Alex Wussow can help with this implementation.

Steps

Step	User Interaction	UI Data	Ux Mockup	Validations	Discussion
	Main success case				
1	Click on "Maintain Script Library"	- Script name - Script description - System name - Target connector - Script type		Supports "create, open, delete, copy" actions - Script type can be "SAP Query," "BW Query" or "Web Service"	See Process Control 3.0 "Rule Script" in "Global Evaluation Setup" where this has been implemented before. Confirmation required: 1. The field System name is not relevant. Should this be removed from the UI? **A: I would suggest we keep it as it is possible we use the Web Service to connect to non-SAP systems.**
2	Select "Create New Script"	- Script ID - Script name - Script description - Script type - System connector (s) - Query name - Org parameter type (e.g. cost center, sales center)		- Script ID is a unique ID# that is auto-generated - Script type can be either SAPQuery, BW Query or Web Service - System connector(s) is a drop down with the available SAP systems. It is restricted based on the script type just selected - Query name is a preview of the name of the assigned query. To select this is to generate a popup to browse and select the actual query - Org parameter is a list of applicable parameters. The applicable parameters are those parameters that are maintained for the organization that are allowable inputs for the query selected. To modify is to select from a dropdown of available parameters. - Return value is the column from the SAPQuery or BW Query that will be returned as the value used for the base KPI in question. To select this is choose from a dropdown the relevant column from the SAPQuery or BW Query	See Risk Management 3.0 "KRI Implementation" under "Risk Monitoring" where all of this has been implemented except for org parameter type and material parameter type. Queries: 1.Could you please add the mockup of query popup here? **I need to go look for this...** 2. Should there be a customization to restrict the SAP queries/BW queries/Web services that would be displayed for the user (like in RM) ? **Yes I think that is the best approach. In fact we will want to pre-deliver a nubmer of queries as part of this customizing.**
2a	Select & assign query (dropdown)	- Return value column - Return value unit - Return value unit type - Aggregation method		- Return value units is the unit of the value returned from the SAPQuery or BW Query. It should be auto-populated based on the return value column selected. Return value units must match the base KPI unit type in order for the script to be assigned to the base KPI in question. - Return value unit type is auto-populated based on return value unit - Aggregation method is a dropdown select of different aggregation options (Arithmatic Average, Number of Values, First Value, Last Value, Maximum, Median, Minimum, Standard Deviation, Sum, Variance) . User must select 1 and only 1 aggregation method.	
2b	Test connector				See Risk Management 3.0 "KRI Implementation" under "Risk Monitoring"
2c	Test script		See above		See Risk Management 3.0 "KRI Implementation" under "Risk Monitoring"
7	Save and Exit				
	Alternative steps				
1	Alternative 1				
2	Alternative 2	Look into OLSP maintenance			
	Failures				
1	Failure 1	- Values not returned for requested organizations			
2	Failure 2	- Values not returned for requestd materials			

Fig. 3. Example wiki page showing integrated use case and UX mockups

Within the outbound product validation work stream, sprints two through eight included co-innovation partner review of the previous sprint's deliverables, collection of feedback, and incorporation of enhancements into the product backlog. As a tool to engage customers in the per-sprint activities, we used the following illustration.

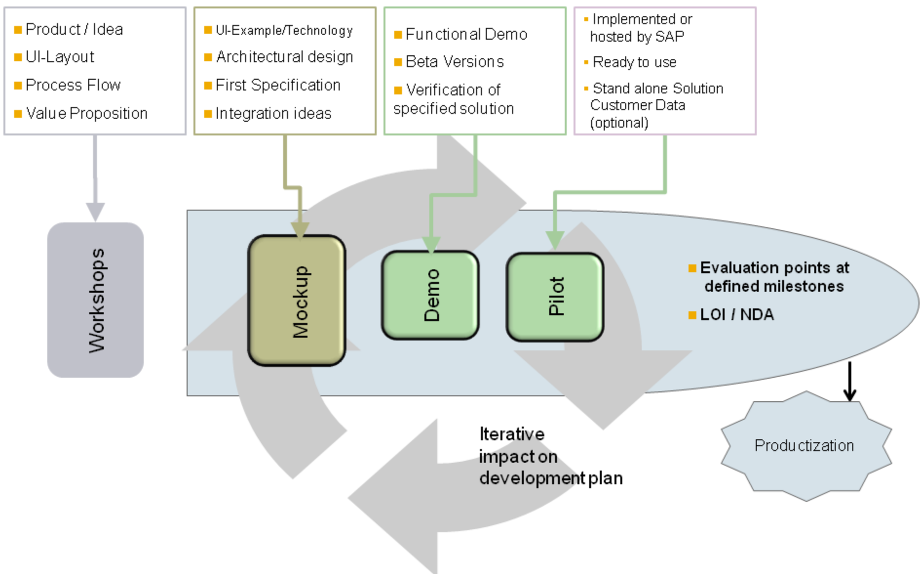


Fig. 4. Iterative customer-focused activities included in co-innovation relationship

Working closely with co-innovation customers throughout development greatly increased the likelihood of customer adoption and veritably ensured a "no-surprises" release. However, coordinating inbound product definition, UX design, development, QA, and outbound product validation activities within short sprints proved to be a sizeable challenge, even with careful planning. An idealized orchestration, defined during the planning phase looked like Fig. 5.

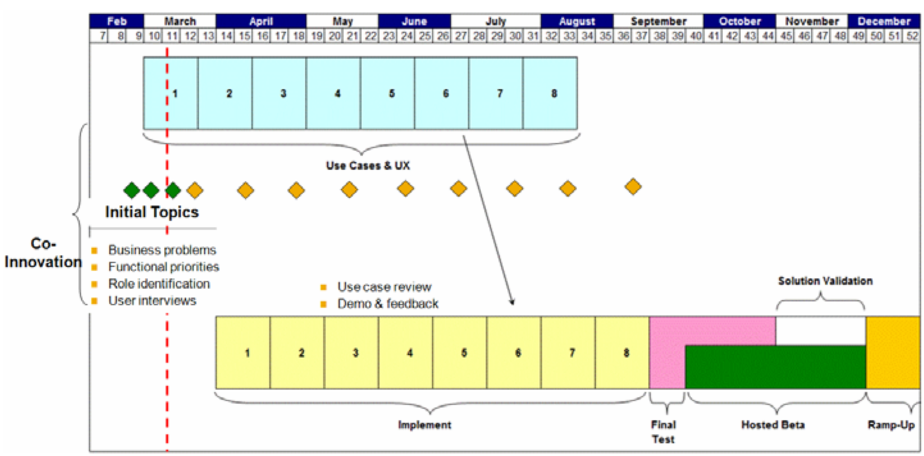


Fig. 5. Draft project plan including UX, co-innovation, development, and testing

In practice, such a plan is rarely as clean as envisioned, and the 1.0 release was no exception. Handoffs from customer to SAP, from Solution Management to UX, and from UX to Development always demanded more energy and commitment than planned, sometimes spilling over well into subsequent sprints. Specific to UX deliverables (wireframes, detailed mockups, interaction flows, etc.), oftentimes ongoing refinements due to technical constraints or fluctuating customer requirements rendered the term “handoff” inappropriate; perhaps “ongoing collaboration” would be a better characterization.

The UX team utilized various techniques to maximize effectiveness and minimize risk in this environment. The most useful and easily transferable to other projects are summarized below:

- Define, refine, and live in an environment of "good enough for now." While many designers’ inclination is to document and deliver only the final, perfected design, Agile Development prefers the “Just Enough Documentation” (JED) approach to deliverables. For UX, JED might mean delivering very rough wireframes, or even taking pictures of drawings on a whiteboard. It will vary per project—the goal is to be flexible and fast, open to the likelihood that changes will be made in future sprints, anyway. “Good enough for now” is often far more effective than “perfect.” However, there are tradeoffs between detail now and detail later. If too little detail is included in initial deliverables, the latter burden of clarifying ambiguities may become overwhelming, distracting from future sprint’s design deliverables. The goal is to find the right balance between speed now and speed later; sometime trial and error is the only tool available in working toward this goal, as each team and project calls for a slightly different definition of “good enough for now.”
- Manage UX dependencies within the development timeline; beware of designing one piece of the puzzle at a time. Agile projects tend to engender a very piecemeal approach to product design, with business and technical representatives eager to cut up features into little bits, prioritize them, and spread them throughout sprints. UX must beware of the risks of designing one piece of the puzzle at a time—the puzzle may not fit together in the end—and address design tasks holistically, considering antecedent and subsequent user interactions. Sometimes, it’s best to request that items on the product backlog be reordered (reprioritized) in order to align antecedent and subsequent interactions.
- Embrace opportunities for incremental Ux improvements within Agile. Sprint reviews, product backlog reviews, and customer engagement sessions all present opportunities to introduce UX improvements to the product; use them wisely. Consider including a devoted “cleanup” sprint at the end of the release to tie up any loose ends and ensure overall UX quality.

4 Back to the User

Wrapping up the 1.0 release and preparing for the 2.0 cycle, SAP contracted User Centric, an independent usability consulting firm, to conduct a formative usability test. Based upon customer ramp-up feedback, productive customer input, and internal review, four tasks for the Sustainability Analyst role were identified as most likely to benefit from usability testing These tasks involved some of the more complex business objects and the most sophisticated interactions:

Task 1	Walkthrough of KPIs
Task 2a	Edit KPI Relationships
Task 2b	Create KPIs from Dialog
Task 2c	Define Formula
Task 3a	Create Core KPI
Task 3b	Define KPI Relationship
Task 4a	Locate Framework
Task 4b	Identify Related KPIs
Task 4c	Add KPI to Framework

Despite the known challenges inherent to the tasks, the report concluded with the assessment: "Overall, test participants viewed the SAP SuPM application favorably." Yet, task completion rates clearly revealed areas demanding further attention:

Table 2. Task completion rates from the formative usability test

Task 1: Walkthrough of KPIs (n=8)			
	Success	Struggled Success	Failure
Number of Participants	5	2	1
	63%	25%	13%

Task 2a: Edit KPI Relationships (n=8)			
	Success	Struggled Success	Failure
Number of Participants	0	4	4
	0%	50%	50%

Task 2b: Create KPIs from Dialog (n=8)			
	Success	Struggled Success	Failure
Number of Participants	4	4	0
	50%	50%	0%

Task 2c: Define Formula (n=8)			
	Success	Struggled Success	Failure
Number of Participants	0	3	5
	0%	38%	63%

Task 3a: Create Core KPI (n=5)			
	Success	Struggled Success	Failure
Number of Participants	4	1	0
	80%	20%	0%

Task 3b: Define KPI Relationship (n=5)			
	Success	Struggled Success	Failure
Number of Participants	4	1	0
	80%	20%	0%

Task 4a: Locate Framework (n=6)			
	Success	Struggled Success	Failure
Number of Participants	6	0	0
	100%	0%	0%

Task 4b: Identify Related KPIs (n=6)			
	Success	Struggled Success	Failure
Number of Participants	5	0	1
	83%	0%	17%

Task 4c: Add KPI to Framework (n=6)			
	Success	Struggled Success	Failure
Number of Participants	2	3	1
	33%	50%	17%

Based upon the report of findings and a secondary prioritization according to expected business benefit, UX prepared approximately twenty design enhancements that were rolled into 2.0, released to the market in late 2010.

Concluding with feedback from a co-innovation customer, the business value of the product is clear:

“*SAP Sustainability Performance Management will enable us to move beyond measuring and reporting to incorporating sustainability in a more strategic way within our business*.” – John Gagel, Manager of Sustainable Practices, Lexmark

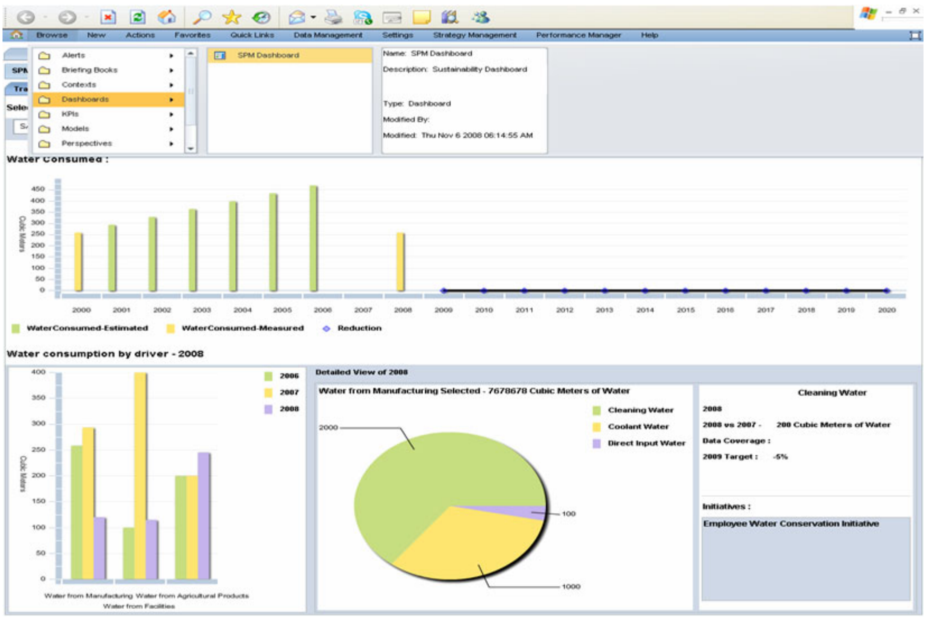


Fig. 6. SAP SuPM example dashboard illustrating historical performance of a water KPI

References

1. Global Reporting Initiative, `http://www.globalreporting.org/Home`
2. Carbon Disclosure Project, `http://www.carbondisclosureproject.com`
3. Dow Jones Sustainability Index, `http://www.sustainability-index.com/`
4. SAP solution overview of Sustainability Performance Management, `http://www.sap.com/solutions/executiveview/sustainability/sustainability-performance/index.epx`
5. SAP’s sustainability report, `http://www.sapsustainabilityreport.com/`

Make User-Centered Design Sustainable in China

JunJie Zhu

SAP Labs China, Shanghai Pudong Software Park, 1001 Chenhui Road,
201203 Shanghai, China
jeff.zhu@Sap.com

Abstract. Sustainability, in a general sense, is the capacity to maintain a certain process or state indefinitely[1]. In this paper, I will discuss 3 phases from survival, growing to sustainable to keep up user-centered design process within organization and develop it in a sustainable way. However, running a design process in the organization level could not just happen without any culture context. This paper proposes the major focuses in these three phases, especially reuse and renewal of user experience design assets internally and externally and illustrates some examples and ideas in context of Chinese culture.

Keywords: Sustainability, design, interaction design, user-centered design.

1 Introduction

Sustainability, in a general sense, is the capacity to maintain a certain process or state indefinitely [1]. In this paper, I claim the sustainability of user-centered design have 3 phases: *survival phases*, *growing phases* and *sustainable phase* in a culture context as shown in *Sustainable Developing Model of User-Centered Design Model* (Fig. 1).

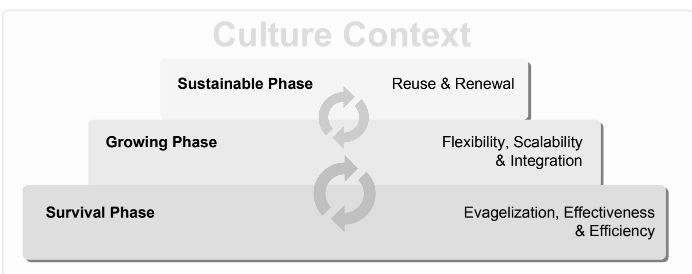


Fig. 1. Sustainable Developing Model of User-Centered Design

In survival phase, I propose audacious evangelization of user-centered design to make all the stakeholders in users' shoes to develop desirable product and service; and focus on the effectiveness and efficiency of the design process to show the design value in a short time frame. In growing phase, *flexibility* and *scalability* could be achieved based on project requirements; and seamless *integration* of user-centered design process with existing product development process as a complimentor could

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 509–518, 2011.
© Springer-Verlag Berlin Heidelberg 2011

extend the design excellence to product development all over the organization with less impact. With the achievements in survival and growing phase as a solid foundation, sustainable phase could be reached to concentrate on *reuse* and *renewal* of assets, such as all the research data, design pattern, the user-centered design methodology and project experience etc. By internal reuse and renewal, all the stakeholders could deliver great product and service in a cheap, deep and fast way. By external reuse and renewal, we are forming a user-centered innovation culture to remove hard use product and system on the earth. However, the whole sustainable development of user-centered design in these three phases is not linear; there are iterations in order to form solid foundation for the higher phase.

2 Survival Phase

To reach a sustainable level, the first thing for design team to do is to maintain user-centered design process in existence. It also can be understood as survival phase, in which all the endeavors are to convince people to initialize the user-centered design process in the organization.

Currently user-centered design is a hot topic in China. The word "UCD" has begun sneaking into the design and business press. Designers think that’s the new way to expand their design capability to have great design. Executives in companies consider it as a critical success factor to have value added product and service with bigger profit margin. With the prospect for better product and service, design team will initialize user-centered design process. In the early phase, there will be less challenge from other departments due to big expectation. As time goes by, as money is spent on the project, the gap between the design and big expectation will cause more and more challenges, especially in the companies with low design literacy. With restrained resource, small and medium companies are more “result” oriented in China. No one will just spend resource on process without outstanding outcome. As a design director in a walkie-talkie manufacturer in the south of China, my friend was once asked by his boss to deliver a successful product design as Apple Ipod with RMB 50,000 (Approx $7,000) budgets in 3 months. He immediately took the user-centered design as a key to try to give solution. That was tough time for him. Design process and methodology can’t guarantee outstanding designs. Actually in this survival phase, design team should help the companies to form a long term view of user-centered design. Evangelization of design thinking and design process along the whole organization is crucial and long-lasting efforts. Once people understand the role of user-centered design, they will have realistic expectation, rather than “gamble” on the design process to have great product in a blind and impatient way.

2.1 Evangelization of User-Centered Design in the Beginning

A several-day long fundamental user centered design workshop could be setup for all the people from different departments. It could be delivered by internal trainers from design team or even external user-centered design gurus. Actually people have their own professional mindsets and speak their own jargons in the product development process. The purpose of this workshop is to spread the seeds of user-centered design fundamentals and break mindsets. Chairman Kun-Hee Lee in Samsung group even went a step further with a long term vision. He established the Institute of Innovation with the help from Gordon Bruce in early nineties. He required all the mid-level managers to attend one year design program, where they learned the design methodology and did design project abroad. All in all, this program incubated the long term design thinking and global view among the mid-level managers in different departments, not just in design department. It formed a solid base to compete with another giant electronics manufacturer Sony. What can we learn from this story? Considering Hofstede's dimension of culture [2], there are two key dimensions of culture, power-distance and uncertainty avoidance, to influence the adoption of a new user-centered design process. In a context of high power-distance and medium to high uncertainty avoidance Chinese culture, Top-down approach will be more effective. Design team really needs an executive level sponsor. With his help, design team can get resource to evangelize user-centered design process and take risk to pilot user-centered design. It takes time and money to influence people within the whole organization to bear user in their mind and always stand in users' shoes along the product development process, especially in an engineering oriented company.

2.2 Focus on Effectiveness and Efficiency

With limited resources in the early survival phase, how can design team concentrate on the effectiveness and efficiency of user-centered design process? Effectiveness, I mean showing the power of user-centered design process in a short period. Efficiency, I mean accelerating research and design speed with limited resources.

Taking account of the budget, timescales, resources, skills and other constraints, a small group of passionate volunteers, not only designers as well as product managers and developers, could work on a limited scope project with high business impact as a pilot project in a short time frame, say 2 to 3 months. Once the pilot project is successfully done, it could be a show-case to evangelize the user-centered design by usability KPI improvement and positive real users' feedbacks on the implementation comparing with old projects. External design consulting firms could also be involved into the pilot project with internal employees. That will also be a good opportunity for internal employees to learn the hands-on experience from user-centered design gurus. Under the sponsorship from a C-level executive, Innovation Design Center in Lenovo worked with a US based design firm to launch a concept project for future computing device. Designers in Lenovo learned a lot in this project. That was the beginning for Lenovo to adopt user-centered design process.

Based on the experience from pilot project, a core set of practical methods could be defined, for example persona, use case etc. In the survival phase, design team don't need the design tool set full of all the cutting-edge design methods in the world, they'll really focus on those best practice design methods on the basis of their applicability, maturity, availability, and cost-effectiveness. These tools are not only design tools, but collaboration tools to convey and document design ideas to all the stakeholders as well. In 2003, a computer manufacturer in the north of China had one designer work for full year to gather all the design methods as much as he can into a thick design method book, somehow like the Method Cards from IDEO. However, at that time, that was the early adoption phase of user-centered design in that company.

Few people, even designers tried those methods out from that book. The outcome of one designer' full year effort was just a thick book. I wouldn't say this is a bad practice to try to form a full design tool box, but rather non-effective and inefficient way especially in the survival phase. Why not just let this designer to initialize the pilot project I mentioned before in this paper? All the stakeholders in the companies would see a successful example rather than a thick design method book that few people will refer to.

Once the successful pilot project is there, limited but effective design methods are defined, another round of evangelization will begin to sustain the visibility of the power of user-centered design to all the stake holders. These continuous user-centered design practice focusing on effectiveness and efficiency, and evangelization of user-centered design best practice and methodology form a positive spiral in survival phase. Top management team will see the business impact from bunch of successful projects and is willing to invest more resource into user centered design practices which will reach another higher level defined as growing phase in the following section.

3 Growing Phase

With the achievement in the survival phase, the value proposition of user-centered design is understood within the whole organization. User-centered design process is established initially and survives with successful projects. However, new design process has a similar life-cycle as product, as mentioned by Geoffrey Moore in “Crossing the Chasm”, starting with a small group of design pioneers, either crossing the chasm to moving into the mainstream to get true acceptance or being failed in the early adoption phase[3]. How to extend the success to cross the chasm to make user-centered design process grow in the product development?

3.1 Expand Successful Activities to Include More People and Projects

It is too easy to see user-centered design as the ‘enemy’ for those people who have failed to serve the needs of users. Such a combative attitude in the company rarely succeeds in the long run. “All the early user-centered design practitioners have been frustrated by the slow acceptance of UCD, even when their organization had clear evidence of its success”, as reported by UPA 99 workshop report[4].

It is not enough to simply repeat the user-centered design activities within design team. There are two dimensions for expansion: train more stakeholders to use successfully piloted design methods and practice in their development work; and pilot more design methods to enlarge the arsenal of user-centered design. As these activities going on, design team can make allies step by step among all the stakeholders, especially those who see and listen to users' pain points on a regular basis from marketing, field service, training and technical support. Design team can work with them to improve the product and help them get credit for finding any problems that design team can fix.

3.2 Achieve Flexibility and Scalability

In the early stage, user-centered design process is just within the design department, as an extra added part to existing product development process. That's why design team just gets limited resources and tight schedule to run user-centered design activities. Don Norman suggests beginning the user research before the project begins in order to gain additional time for user research in the beginning phase. Under this situation, flexibility and scalability of the design process are the key drivers.

Flexibility, I mean effective personalization of design toolset to meet key requirements from projects. Just as flexible manufacturing system (FMS), the product line can be tailored according to different products manufactured; design team also can pick the right design methods and tools to ensure effectiveness and efficiency. For example, Expert review and formative usability testing are both good choice for design evaluation, the decisions to go for which method or even both methods depend on the project budget, time and usability maturity of the company.

Scalability, I mean running user-centered design process in full scale or partial scale based on the project need. In SAP Analytics Customer Validation Project, "due to the fact that this project mainly dealt with the validation of already developed Analytic Applications and a tight project schedule, the project team decided it had to skip early user research"[5]. However, it was still a user-centered design success story in Sapdesignguild.com. If the goal of project is to improve current product according to users' feedback, which is known as incremental innovation, we can do less early user research even omit the early user research part, because we have domain knowledge within product managers and designers, because we have users' feedback from sales and support team, because we have early site visit reports and usability testing report of the current product. If the goal of project is to create a new product category in the market, known as radical innovation, then more money and time will be spent on the early user research to identify the product opportunities and more design iterations from users' feedback are required. In this case, there is no doubt to run the project in a full user-centered design process.

A guideline to achieve flexibility and scalability could be provided internally. Project team can follow it to decide the selection of methods and the scale of user-centered process. However, project team should make their own decision based on the experience. Rigid alliance with guideline is not suggested.

3.3 Integrate UCD as Complimentor to Product Development Process

The integration of UCD will happen in the whole organization wise. Definitely it will change people's work territory and power influence. Responsibilities of some departments are reinforced and some are weaken. In engineering oriented company, engineers and developers treat product development as their own responsibilities. A "Bottom Up" way of process integration is tough. Design team can't only drive it from bottom by itself. A "Top Down" approach will be more effective. Executive level officers could drive the integration in a long term and company wise view with the help from design department. The goal of integration is not to compete with or replace current existing product development process, but rather to optimize current product development process into a user-centered manner.

Decouple user-centered design and usability. People have mindset about user-centered design to limit it within usability and design. However user-centered design promotes all "development proceeds with the user as the center of focus"[6]. It is not just a way to design, it is about how to deliver technology feasible, business viable and user desirable products providing holistic user experience along the product lifecycle from creation to disposal or reuse. Usability is just one aspect of the holistic user experience. Designers need to take account of the holistic user experience, not just the ownership of the “usability”. However, all the stakeholders should work together to deliver great holistic user experience. Adobe has a program called developer’s day, in each month, developers have one day to spend on site visit to collaborate with end users with designers and user researchers. Oracle launched the user-centered development process to ensure all the stakeholders share the same vision and user-centered activities. Cognetics Corporation defined a framework for user-centered design called LUCID, for Logical User-Centered Interaction Design. It includes “Support” stage as the last stage, where the designers may work with the technical support group to anticipate and solve design problems[7]. User centered design will really expand all the stakeholders’ vision to think big. People will not only stay in their own territory and focus on the very technical details, but also understand the strategic value of their efforts to make great product and service.

Lower the impact to current product development process. Any user-centered design activities must fit into a well-established product development process in place. There may be no formal methodology or new activities with little impact on the existing process. It is important to identify who will be affected by any changes in the process which is proposed. In the era full of change, Chinese people have readiness to change if it could bring opportunities. With the big vision for great product and service, people affected by the process change could work with design team cooperatively to find new opportunities.

Integrate design methodology, design tools into the new process. In this phase, the arsenal of design methods in design team is keeping enlarging. During the process integration, there is opportunity to embed some successful user-centered methods into product development process. Just as use case, once it is define in formal development process, it plays a key role within solution managers, designers and developers in SAP. Officialization of these design methods in the product development process would influence the way people use them. In the thick book of design methods story in the Section 2, why no one refer to it? Because Chinese people incline to authorized stuff, those design methods are not officially defined in the process. If there are resources, internal design tools could be developed to include all key design methods. For example, a database of user research reports, or a development tool to translate high-fidelity prototype seamlessly into implementation smoothly such as Microsoft Expression.

4 Sustainable Phase

In the survival and growing phase, user-centered design will get enough credibility by encompassing the entire user experience. Project experience and knowledge about user-centered design and trust from all the stakeholders form the solid foundation to reach sustainable phase. In the sustainable phase, the focus will be how to keep user-centered design developing continuously in a sustainable manner.

4.1 Make User-Centered Design Cheaper, Deeper and Faster

The challenge for companies that want to grow into unfamiliar markets is not only understanding the new users and cultures, but also doing so quickly. In a completive commercial world, there are always pressures to deliver great products more rapidly and more cheaply. The mindset of fast, cheap, good – pick two, you can't have all three, have been there for years. Is it an impossible mission to make user-centered design cheaper, deeper and faster? We still have opportunity from following ways.

Design pattern. "A pattern is a collection of elements and their relationships which can be repetitively reached or used in analysis, design, development and use (of cooperative systems)"[8].Based on sufficient knowledge about users and design, a large design pattern library could be formed to support enough design flexibility by user-centered design approach, which means every design pattern will provide solution for frequently-occurring problems and situation based on user research and evaluation to ensure high feasibility, viability and desirability. Designers, product managers and developers can quickly configure standardized design patterns to complete the product design to meet users' needs, which ensure them to focus on "What to be designed" rather than "How to be designed". They can spend more time to validate the whole concept in the product offering level, rather than drill in detail to validate usability of one drop-down box.

Design pattern library is not fixed; it is continuously growing with projects. If there are any valuable user feedbacks to the design patterns in the project, quick improvement could be made to design patterns and central update of the design pattern library keeps all the design using these patterns up-to-date with continuous improvement. As an example, in the different validation projects, we found a lot of end users complaining about table row selection in SAP Business by Design applications. Once quick improvements were implemented into the design pattern based on this feedback, this issues are gone in all the applications. Along with improvement of current design pattern, specific design can also be transformed into design pattern if it proves its value for generic use.

Design guidelines. Design guideline is a way to compile successful user-centered design experience to guide the ongoing product development. It can ensure design consistency, not only from look and feel level, but also from holistic user experience level. Designers, product managers and developers in charge of different products can quickly refer to it to find existing solution to certain issues. Design guideline could be a tool in effective collaboration among all the stakeholders. Designers, product managers and developers won't waste time to stick in the issues which have been solved in other projects in design guidelines. Some companies even put design guideline in the Wiki so that all the stakeholders can ask questions and post successful design experience and solution in the Wiki.

User relation management. Customer Relation Management is about how to maintain and support current and potential customers to sale product and service. User relation management in my words means to keep in touch with some key end users for participatory design in a long term way.

It is time-consuming to recruit qualified end users for user research. Based on my experience, even we did a lot of homework with recruiting company in validation project, the hit-rate of qualified end users will be around 2 to 3 for SAP applications in Shanghai. The hit-rate is even lower in US and Germany. Lead user is treasure for user research. Thom Hogan is an encomiastic Nikon fan. He writes review of every Nikon professional camera and user guide of those cameras in eBook format for sale. In his website, you can find his prediction and comments about Nikon camera and even the way how to do marketing with the latest camera. Bjørn Rørslett writes review about almost every Nikon lens ever made from sixties. Nikon often sends pre-production camera model and lens to them for review. Their insights to Nikon camera were constantly adopted by Nikon. It is lucky for Nikon to have loyal users to improve product and service with them.

In this win-win case, end users could try latest product and companies get valuable feedback in return. All the stakeholders in the product development should think a way to maintain long term relationship with those key end users, such as providing early try-out of products, free admission to technology conference, training, books and incentives. Recruiting companies always consider the list of users they find for companies is their own assets. It is valuable assets for company too. I once worked on a project with recruiting company, which tried to keep some capable end users for long term product evaluation, not just invite them once in usability testing.

Recruiting users from agency is an expensive way, but not the only way. Online technology and product forum is gold mine full of capable lead users. Personal networks and online social networks can ensure us to reach the right users we want. And they are low cost ways. Especially in China, with a high context culture, searching users from personal networks is the most effective way. Actually Chinese people like to build Guānxi, which describes the basic dynamic in the complex nature of personalized networks of influence and social relationships, and is a central concept in Chinese society[9]. I once did a project called "on/off" with researchers from Institute of Design in Shanghai, we did shadowing observation and interview with a family with 5 people in one week. When project was over, we became friends with the big family, they even invited us for a dinner to show hospitality. It won't be tough for researchers to build relationship with end users in China. As projects going on, a handful of qualified end users will be available in the database. New project team can quickly search qualified users in a few mouse clicks. Design team in ChangHong, one of the largest Chinese consumer electronics manufacturers, is going on this direction with a user database containing 1000 users in different cities in China to support design activities[10].

User research tools. If we go further with that idea of database of end users, we even can make a database of design research, which means it contains not only user profile, but also their activities with product and service, their comments and their insights to the product and service. Patrick Whitney and Vijay Kumar proposed activity based research to quickly make surprising discoveries about product opportunities in a new market. They "achieve this neither by looking at the specific product nor by looking at the general culture, but by focusing on people's activities when they are using a product or service a company wants to develop."[11] I worked with Patrick Whitney in the "Global company in local market" project to develop software tools for managing research data from video ethnography and disposable camera study. Theses software tools "are used for tagging the video clips as representative of certain activities, making comments about the activities, recognizing patterns of people's behaviors, and extracting insights to help in the development of new products and systems"[11]. It is expensive to run this kind of database on an ad hoc basis, however it will be much quicker and cheaper way, considering reuse of these qualitative research data for relevant projects.

Distributed co-creation. "Distributed co-creation, the production of joint work products by loosely coupled networks of participants using iterative and decentralized approaches, is becoming an increasingly important mode of innovation" [12]. It becomes a trend for software companies to build an idea platform to motivate customers to summit their idea to improve current product design. To name a few, Google has its Google Product Idea, Dell has its IdeaStorm. Idea Scale and User Voice are also providing services to companies with pre-made idea platform to help them to roll in product ideas. Those idea platforms, no matter homemade or from service vendor, are instant and low-cost channels for software makers to reach numerous real end users with their brilliant new ideas. End users' will to participate into these idea platforms is really high, since they are interested in making a contribution and seeing it implemented.

Design pattern, design guideline and user relation management will ensure development efficiency and consistent user experience in the context of reuse and renewal. This will also charge less for customers in the final product in turn, because less time and less resource is spent on the product development.

4.2 Promoting Maximal Renewal and Reuse

In the early phase, company tends to hide their user-centered design process as a secret weapon for innovation especially in a high context culture. With a design literate perspective, user-centered design process is just like a mathematical formula, capable person can solve difficult mathematic problem by it, however incapable people can't. In a wired world, sharing of knowledge is win-win case for all the parties. SAP, IBM and other companies all share their user-centered design process in publications and websites. People can quickly find the details about SAP user-centered design process in www.sapdesignguild.com. Apple and Microsoft share their design guideline of operating system. Yahoo shares its design patterns. Those assets are no longer internal, they are external. People and companies learn each other to keep the best practice of user-centered design evolving. I once interviewed a designer candidate in SAP, he told me he kept reading articles from www.sapdesignguild.com to reinforce his knowledge about user-centered design. There is no doubt, when he became my colleague, he caught up the design work quickly without dedicate training. That is an amazing example how knowledge sharing educate future employee as a reward. In this level, renewal and reuse of design process and methodology is achieved not only within the company, but also within the industry even across the industries. User profile database and user activity database discussed before can also be shared among companies in this way. As we know, user research is costly and time consuming. With joint user research activities among companies and universities all over the world, these databases can be reused and renewed constantly for design and education purpose with cost sharing. Companies can extract different insights from users and those databases based on their own design capability in a much quicker, cheaper way. SAP co-founder Hasso Plattner invested in D-school in Stanford University and Hasso Plattner Institute for Software Systems Engineering (HPI) in Potsdam, Germany. This is a step further than Chairman Lee in Samsung I mentioned in Section 1. These two schools focus on design thinking to foster a user-centered innovation culture, no matter the students are SAP employee or not. This is another way to educate future employee with user-centered mind. A company full of people with user-centered mind will be no doubt a user-centered company. In the ideal situation, if all companies are user-centered company, those hard-to-use product and system will no longer exist on the earth, which is effective means to reach the ultimate ecology to utilize the limited resource.

References

1. Wikipedia, `http://en.wikipedia.org/wiki/Sustainability`
2. Hofstede, G.: Culture's Consequences. Sage Publications, California (2001)
3. Moore, G.A., McKenna, R.: Introduction: Crossing the Chasm Marketing and Selling High Tech Products to Mainstream Customers. HarperBusiness, New York (1999)
4. Quesenbery, W.: UPA 99 Workshop Report: Crossing the Chasm - Promoting Usability in the Software Development Community. UPA Common Ground 10(1) (2000)
5. Weissenberger, U.: User-Centered Design Success Story: SAP Analytics Customer Validation Project. SAP Design Guild (2006)
6. Rubin, J.: Handbook of Usability Testing:How to Plan, Design, and Conduct Effective Tests. John Wiley and Sons, Inc., Chichester (1984)
7. The LUCID Framework, `http://www.cognetics.com/lucid/`
8. Alexander, C.: A Pattern Language: Towns, Buildings, Construction. Oxford University Press, New York (1977)
9. Wikipedia, `http://en.wikipedia.org/wiki/Guanxi`
10. User-Centered Design Practise in ChangHong, UPA 2007 Presentation, Beijing (2007)
11. Whitney, P., Kumar, V.: Faster, cheaper, deeper user research. Design Management Journal, Fusing Design, Strategy and Technology 14(2) (2003)
12. Bughin, J., Johnson, B., Miller, A.: Distributed Co-Creation: The next wave in innovation. McKinsey&Company (2007)

How to Improve User Experience in Mobile Social Networking: A User-Centered Study with Turkish Mobile Social Network Site Users*

Özgürol Öztürk and Kerem Rızvanoğlu

Galatasaray University, Faculty of Communication,
Ciragan Cad. No:36 Ortakoy 34357 Istanbul, Turkey
{ozozturk,krizvanoglu}@gsu.edu.tr

Abstract. This study evaluated the hypothesis that design aspects of the user interface of the applications and the capabilities and constraints of different mobile devices could affect mobile user experience in social network sites (SNS). For this purpose a qualitative mobile usability test, based on a multi-method approach, was carried out with a sample of 25 Turkish university students who were experienced mobile SNS users. The tests were conducted with two different smartphones: Iphone and Blackberry. The participants were observed during the task executions and additional data was collected by the "think aloud" procedure, eye-tracking and video recording of the participants. A short debriefing interview was also made to gain a detailed insight into the user experience. The findings revealed significant mobile usability problems caused specifically by the user interface features of the applications and limitations of each device and enable to propose guidelines to improve user experience in mobile SNS.

Keywords: User Experience, Mobile, Social Networking.

1 Introduction

With the evolution of the mobile platform and the rapid adoption of mobile devices such as smart phones, social networks, which began as web-based applications, have migrated onto the mobile platform [1]. As social networking sites (SNS) can easily provide a base for self-expression and positive experiences of social pleasure, they are widely used on both platforms. Although, social networking ranks as the fastest-growing mobile content category with huge numbers of users, user-centered studies contrarily showed that SNS mostly failed to improve user experience in both platforms. The studies point out the need for employing diverse methods of evaluation in further studies to understand the holistic user experience in SNS [2] .

The purpose of this study is to explore the Turkish user experience in mobile SNS and generate guidelines to improve mobile usability in SNS. With a young population and a high rank in both mobile device ownership and SNS usage, Turkey is an important case for user-centered studies on mobile usability in social networking.

---

* This study was realized under the coordination of Assoc. Prof. Kerem Rızvanoğlu with the support of Galatasaray University Scientific Research Fund.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 521–530, 2011.
© Springer-Verlag Berlin Heidelberg 2011

This study, which is part of a larger holistic study on social networking among university students in Turkey, evaluated the hypothesis that the interface design aspects of the applications and characteristics of the mobile devices could affect mobile usability in SNS. For this purpose a qualitative mobile usability test on Facebook, based on a multi-method approach, was carried out with a sample of 25 Turkish university students who are experienced mobile SNS users.

2 Theoretical Background

Social networking sites (SNS) are one of the major phenomena of Web in recent years. Boyd and Ellison [3] define SNS as web-based services that allow users to share a public or private profile with common users and explore connections with others within the site. In the last few years, the immense interest of the users towards the SNS brought the emergence of various studies on the usability of this phenomenon [4].

Mobile social network sites (MSNS) enable groups of friends to be accessed and engaged with from one's mobile phone. Much like social network sites on the Internet [5], [6]; these services may help users to build valuable networks through which to share information and resources [7]. Mobile social networks are the impetus for the creation of an entirely new sub-industry in the wireless sector, thus representing a new aspect of wireless innovation, and increasingly are providing a platform for content and technological innovation in the business environment [1].

Only a limited number of studies investigate how people integrate MSNS in their daily lives. Humphreys [7] examined the social and behavioral norms of Dodgeball, which is a MSNS that seeked to facilitate social connection and coordination among friends in urban public spaces. The findings of that study suggested that Dodgeball use could influence the way that informants experience public space and social relations therein.

User experience in mobile platforms appears to be one of the most promising topics of research in the area of human-computer interaction, as people suffer from usability issues caused by the design of mobile interfaces and limitations of the mobile devices.

Forrester report [8] showed that Internet on mobiles suffered from three major problems today: 1) Mobile tailored content (mobile versions of the web sites that are accessed by a browser on a mobile device) is hard to find. 2) Usability is poor: Typing with mobiles is difficult, mobile tailored sites are not properly designed for mobile use, but rather modified slightly from Internet offering. There is also lack of consistency between the sites and within the site. The downloadable applications provide better usability, but the applications are hard to find. 3) Access to Internet on mobiles costs a lot and users do not understand how the cost is generated. Steven Browne's (2007 in [8]) findings overlapped with those of Forrester report. His findings supported the notion that Internet on mobiles suffered from three major barriers in user friendliness: Mobile search is inaccurate, carrier / operator portals are ambiguous and input is cumbersome.

Luke Wroblevski [9] is a pioneer of "MobileFirst" paradigm and believes that designing for mobile first can open up new opportunities for growth and lead to a better overall user experience. In this context, he identified the constraints and capabilities of mobile devices. According to Wroblevski [9], the constraints are screen size, slow connections and context of use whereas mobile device capabilities are multi-touch input from one or more simultaneous gestures; precise location information from GPS / cell towers / Wi-Fi; user orientation from a digital compass; device positioning from an accelerometer; and integrated audio, video and photo input.

Nielsen was one of the experts, which insistently emphasized the need to improve mobile user experience. Nielsen [10] claimed that -in terms of user experience quality- mobile web could only provide an experience that was observed in desktop web environments back in 1998. Norman and Nielsen [11] also focused on the mobile user experience through devices that provide gestural interfaces. Norman and Nielsen [11] investigated the user experience problems in gestural interfaces by referring to fundamental principles of interaction design and claimed that these interfaces revealed a step backward in usability. They proposed several reasons for this assumption: The lack of established guidelines for gestural control, the misguided insistence by companies to ignore established conventions and to establish ill-conceived new ones and finally the developer community's apparent ignorance of the long history and many findings of HCI research.

3 Methodology

The purpose of this study is to explore the Turkish user experience in mobile SNS and generate guidelines to improve mobile usability in SNS. Below are the two research questions of the study:

Research Question 1: How do design aspects of the graphical user interface of mobile SNS applications affect mobile user experience in SNS?

Research Question 2: How do capabilities and constraints of different mobile devices affect mobile user experience in SNS?

This qualitative study was based on a multi-method approach, which consisted of a background questionnaire, task observation and a structured debriefing interview. The background questionnaire provided information on demographics, mobile web and mobile social networking experience of the participants. The study also employed observation methods of data collection in order to gain better insight in mobile social networking. The tests were conducted with two different smartphones which constituted the most popular platforms in Turkey: Iphone and Blackberry. The most popular SNS in Turkey, which is Facebook, was chosen for the study and two different tasks were designed for the users to perform with Facebook applications peculiar to each device. The users were asked to terminate each task in 6 minutes. First task was to "log in Facebook, update the status and send a message to a friend". The second task was to "upload a specific picture in the gallery, tag it and write a comment on it". All the participants who finished these tasks were asked to "log out". In this context, a sample of 25 Turkish university students who were experienced mobile SNS users were involved in the study. The sample included 13 IPhone users (7 male, 6 female) and 12 Blackberry users (6 female, 6 male).

The tests were conducted in "7th Sense Customer Insights Lab" which is a full-equipped usability laboratory established by Turkcell in İstanbul. The navigation was directly observed and recorded on a structured observation sheet by the researchers. Besides the observation, additional data was collected through a head-mounted eye-tracker and video recording of the participants. Final structured debriefing interview provided complementary findings.

The analysis framework was derived from the study of Wroblevski [9] and Nielsen and Norman [11]. For the discussion addressing the first research question, the capabilities and constraints of mobile devices in [9] were referred. These were integrated with the fundamental principles of interaction design –that are completely independent of technology- mentioned in Nielsen and Norman's [11] study. The analysis framework that was derived from these two studies include the following parameters: Visibility and Scalability, Consistency and Standards, Discoverability and Feedback, Reliability, User Diversity, Bandwidth and Speed in Mobile Network Performance, Awareness and Use of the Capabilities in Mobile Phones.

4 Results and Discussion

The findings from the background questionnaire confirmed that participants were experienced users of mobile SNS. At the task-observation stage, it was observed that most of the participants achieved to execute all the tasks. As an exception, most of the Blackberry users failed to "log out". From 12 participants using Blackberry, the number of participants who failed to execute specific tasks is as follows: Status update (n=3), tagging (n=2), log out (n=10). From 13 participants using Iphone, the number of participants who failed to execute specific tasks is as follows: Status update (n=2), tagging (n=2), log out (n=2). Task execution periods were close for each group. Blackberry users finished the first task in an average of 3,6 min and the second task in average of 5,5 minutes. Iphone users finished the first task in an average of 3,7 min and the second task in average of 4,8 minutes.

4.1 Visibility and Scalability

One of the fundamental principals of interaction design is visibility. It is also called "perceived affordances" or "signifiers" [11]. It was observed that participants tend to compare their experience in the application with their mobile or desktop browser experience and chose to use the app suboptimally. Thus the major complaint was "*I can not see and control all of the functions in one page, I have to find them and it takes too much time. Applications' logic is different from Facebook on browser*" (Eİ06). Even if the participants could see the buttons, some couldn't understand the function of the icons, as there were no labels accompanying them. In order to maximize the number of functions in the user interface, the labels were avoided.

User frustration caused by the violation of the visibility principle was also observed when the participants were asked to "log out" from the app. The navigation strategy was based on browser experience. The participants either attempted to scroll down to the bottom of the profile page, which was the typical location of "log out" of the SNS on mobile browser, or they oriented to the header area in the app, as it was the typical location of "log out" on a desktop browser. In Blackberry SNS app the "log out" function was hidden deep under the "options" button, which could be accessed through the "menu" button. Inevitably, most of the Blackberry users failed to achieve the task of "log out".

There is a plethora of screen sizes for cell phones [11] and the visibility in the screens of mobile devices is fairly limited when compared to PCs. There simply isn't room for any interface debris or content of questionable value [9]. It was observed that small targets that work well with QWERTY keyboard and navigation ball in Blackberry were inappropriate for fingers on the touch screen of Iphone. Although larger screen of Iphone provided more visibility, it had its own problems with control sizes, which caused accidental activation of actions. In this context, some of the Iphone users were observed to touch things by mistake or make a gesture that unexpectedly initiated a feature.

4.2 Consistency and Standards

Our findings support the notion that "consistency and standards" appear to be the most problematic issues in mobile Facebook application user experience. User experience on Facebook through desktop browsers is fairly stable. In contrast, applications on each mobile device provided completely varying user experiences.

Firstly, most of the participants complained about the lack of functions, which they regularly used on desktop or mobile browsers. In this context lack of chat and games were mentioned as examples. Unfamiliarity with a browser's user interface limited the user's options. Just because of the familiarity with a desktop browser interface, most of the participants preferred to use mobile browsers for accessing to Facebook. Although on mobile devices, applications are easier to use than websites, the obligation to search, understand and learn the core functions on the apps demotivate and led the users towards mobile browsers. Having difficulty in associating the app interface with that of the website, the participants were not able to find the relevant functions for the execution of some tasks.

Tagging a photo seemed to be a major problem on both devices. On Blackberry, the only way to tag a picture is to do it before the upload. Missing that step results in the failure of the participant. However, desktop browser experience enables tagging after the upload and this standard inevitably contradicts with the interaction proposed on Blackberry app. On the other hand, Iphone users also had troubles in tagging a photo since the gestural interface in Iphone Facebook app also violated consistency and standards of desktop browsing experience. Assuming that clicking on the picture will enable tagging, participants tried that gesture to achieve the task. However, despite the fact that some of the Iphone applications enabled to interact with the picture directly and implied the same opportunity for every app, Iphone Facebook app did not support this interaction. This limitation led the users to a random trial-and-error strategy, which included clicking on any button available. The eye tracking data revealed that most of the users looked at the icon at the bottom left corner of the screen which was actually used to tag and like, but they didn't click it until they tried out all other opportunities, because it was very similar to the 'share/send picture' icon of the operating system, a twisted arrow in a square (Fig. 1).

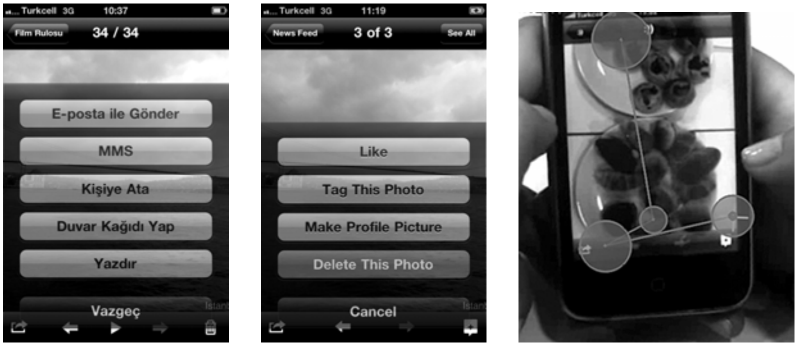


Fig. 1. Although the icons at the bottom left corner have different functions they are identical (*Left, camera roll of the device; middle, interface of the application; right, eye-tracking screenshot*)

The tendency to scroll down to the bottom of the page for finding a logout option was also determined due to the current mobile browser habits. As a final observation, the similarity of desktop and mobile browsers in terms of consistency and standards also seemed to affect user choices for specific actions such as delete.

4.3 Discoverability and Feedback

The true advantage of the Graphical User Interface (GUI) is that every possible action in the interface could be discovered through systematic exploration of the menus [11]. However, our findings showed that this notion could only be useful if it was supported with feedback. Limitation of the "discoverability" led the participants to the adoption of an unstructured navigation based on a trial-and-error strategy, which resulted in errors and abandoning the task.

In addition to that, regardless of the device, most of the participants were not sure about their actions during "photo uploading", "tagging" and "writing a comment". They complained about the lack of system feedback while uploading a photo, which resulted mostly in the re-execution of the task. They could only be sure about the procedure when they saw the photo on their wall. After the tagging it was difficult for Iphone users to find the current tags on the picture, as there were no indicators visible. The task of "sending a message to a friend" also revealed the frustration caused by the lack of relevant feedback as the app didn't provide any notification about the execution of the task. It was clear that the result of an action should be provided as a feedback so that multiple inputs or endless efforts for a "prevented but not indicated" action such as video uploading could be precluded.

4.4 Reliability

A basic foundation of usability is that errors are not the user's fault; they are the system's (or designer's) fault for making it too easy to commit the error [11]. However, lack of "consistency with previous desktop browsing experience", visibility and feedback also brought reliability of the app into question in mobile social networking experience. As in the case of "tagging a photo" or "logging out" in Blackberry Facebook app, by being forced into a random trial-and-error approach rather than an intuitive navigation, the participants lost their sense of controlling the system and did not understand the connection between their actions and results.

4.5 User Diversity

User diversity is one of the core issues that should be considered in designing apps for mobile devices. Especially users from different cultures have specific needs in interacting with user interfaces. Language is one of the key issues in developing localized versions of mobile apps. Our study provided findings to support this statement.

It was observed that the participants didn't understand the "status update" task, which was addressed to them in Turkish with the exact phrase written on Turkish version of Facebook: "*Durum güncelleme*". However, though it was a direct translation of the English phrase, it did not match the local way of expressing the specific action. It was observed that users generally understood "update" as "refresh" and committed the relevant action. When the users were asked for an alternative term to define this specific task in the debriefing interviews, various suggestions were made: "About yourself", "What are you thinking now?", "Share status", "Share on the wall", "Write on the wall", "Share location (with an @)", "Write a message to yourself", "Write a comment on the status", "Update the homepage" and etc. The suggested expressions revealed that the participants found the emphasis "to share" more reasonable to define this action.

The second fundamental problem "lost in translation" was observed during the "log in and out" procedure. Some of the participants had difficulty in writing Turkish characters (*ıüöğçş*) for "log in", although the device had the capability to input these characters. The same confusion and related failure was also observed for the "log out" procedure. Especially in Blackberry, the menu button led to list of commands concerning the app. The last two commands in this list were related with the activity stacks of the operating system (OS) that was in Turkish. The last command in this list was the Turkish version of the word "close". Actually this command enabled to exit from the application. The participants who randomly tried to find the "log out" associated this action with this "close" command. It was even observed that although some of the participants viewed the "options" command which actually provided direct access to "log out", they were attracted by this Turkish command that meant "close" and preferred to use it for the demanded task. It is evident that the language differences between the OS and the app generated a tremendous confusion and failure in user experience.

It was also observed that each cultural group adopts specific strategies and behaviors in mobile social networking. "Messaging" revealed behaviors peculiar to Turkish users. It was observed that Turkish users preferred messaging for various reasons. The most common type of messaging was writing on the wall of a friend where the user expressed his opinion in public space. Although some users preferred direct messages addressing the inbox of the receiver for private purposes, most of the users preferred to send SMS for that case. It was also found out that the users could employ pragmatic strategies through mobile SNS. If the users could not make direct calls from their phone because of financial reasons, they preferred to send messages by SNS through wireless connection to inform a friend about this situation.

4.6 Bandwidth and Speed in Mobile Network Performance

In the context of bandwidth, there seemed to be both positive and negative aspects that affected the mobile user experience on SNS. “Sharing instantly” was evaluated as one of the most popular peculiarities of the mobile devices and an integrated camera enabled users to share pictures anywhere, any time. The photo upload speed of the application was also considered to be faster than the mobile browser. However, the application is not capable of uploading multiple photos whereas the mobile or desktop browser is. Besides, though it is possible to take videos with a mobile phone, it is not possible to upload them through mobile connection. The participants declared that the streaming or downloading speed of a video, which was fairly low, could easily cause frustration, which resulted in the rejection of watching video on mobile. It was evident that the participants didn’t even tolerate small delays and always chose the shortest path in executing a particular task. When they were asked to “write a message to a friend”, instead of finding them from a list, most of them preferred either to reply someone from the inbox or write someone whom they could see his activities on their wall / newsfeed.

In the context of bandwidth and speed in mobile network performance, the findings showed that users preferred desktop networking to mobile browser. The application appeared to be the last choice compared to the latter.

4.7 Awareness and Use of the Capabilities in Mobile Phones

Although mobile capabilities could open up different ways of thinking about interactions between people, data and their immediate surroundings [9], our findings showed that the participants were neither aware of them nor preferred to use them as they weren’t satisfied with the quality of services these functions provide.

Certain number of users stated that the specifications of the device camera were not sufficient so they didn’t use it for taking pictures unless it was very urgent. Even in that case some tend to transfer the picture onto a PC, edit it first and then upload it.

Although there is a function called ‘Places’ in Facebook that is based on location-awareness capability of the device and enables the share of the location information, it came out that none of the participants have ever used it and most of them even didn’t know what it was for. However, for location input they had their own strategies. They just put an “@” sign and wrote the name of the location after it on their wall as a status update.

Even if Iphone allows multiple orientation and gestural controls; only two users used horizontal orientation during text input and stated that the touch targets were small for their fingers in vertical position. It was observed that the participants were experienced users of other gestural controls like pinching, double clicking and sliding, but the application didn’t offer these novel interactions to support mobile user experience in SNS.

5 Conclusion

Mobile social networking sites are the fastest growing mobile content with huge numbers of users, but they fail in improving user experience. This study supported the notion that the design aspects of the user interface of the applications and the capabilities and constraints of different mobile devices may affect mobile user experience in SNS. Considering the problems observed through user experience in Facebook applications of two major smartphones (Blackberry and Iphone) by referring to the analysis framework, the study provided the following implications for the improvement of user experience in MSNS:

- Value visibility. Prioritize the most important set of features for your application, determine the key tasks and locate them with a visual hierarchy such that they provide enough scalability for various screen sizes in mobile phones. Provoke predictability with clear graphics and do not avoid labels for a clearer layout.
- Enable systematic exploration of user interface features and increase discoverability through perceived affordances. Do not hide important functions deep within the menus without signifiers.
- Stick to consistency and solid standards derived from basics of interaction design before the specific interface guidelines provided by different OS developers. Lean on the familiarity with the desktop browser, start with a mobile companion and head for a unique and better mobile user experience. The combination of mobile and desktop experiences could result in more engaged users across both sets of devices [9].
- Provide reliability in mobile user experience based on an intuitive navigation. Prevent accidental activation and triggering of actions through touchscreens in gestural interfaces by providing optimal target sizes. Employ gestures consistently across apps and rely on generic commands [11]. Support sense of control and awareness about the connection between the actions and results of the users by visible signifiers.
- Know your users. Step out of the designer's mental model and be aware of the needs of the target users such as language. Prioritize localization and do not limit it with plain translation. Speak the users' language and provide consistency for language at both the OS and the application.
- Consider the bandwidth of your target users and support technologies that provide a faster mobile network experience anytime, anywhere.
- Develop applications that fully benefit from novel interactions enabled by various mobile capabilities like multi-touch input, location awareness, etc. However, do not forget that these interaction styles are still in their infancy, so it is only natural to expect that a great deal of exploration and study still needs to be done [11].

Considering the lack of user-centered studies on MSNS experience specifically in Turkey, this study contributed to the relevant literature by providing findings to improve user experience in mobile applications. In order to a gain more insight on various aspects of mobile user experience, further empirical studies with divers users in larger groups are needed to be conducted.

References

1. Ziv, N.D., Mulloth, B.: An Exploration on Mobile Social Networking: Dodgeball as a Case in Point. In: International Conference on Mobile Business, pp. 21–22 (2006)
2. Hart, J.: Exploring the Facebook Experience: A New Approach to Usability. In: Interfaces, vol. 77. British HCI Group (2008)
3. Boyd, D.M., Ellison, N.B.: Social Network Sites: Definition, History, and Scholarship. J. Computer-Mediated Communication 13(1) (2007), http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html
4. Ho, V.: Applications of Usability Principles on a Social Network (2009), http://creativebriefing.com/applications-of-usability-principles-on-a-social-network/
5. Boyd, D.: Friendster and publicly articulated social networking. In: CHI Conference on Human Factors in Computing Systems, pp. 1279–1282. ACM Press, Vienna (2004)
6. Ellison, N., Steinfield, C., Lampe, C.: The Benefits of Facebook "friends:" Exploring the Relationship between College Students' Use of Online Social Networks and Social Capital. J. Computer - Mediated Communication 12(3) (2007), http://jcmc.indiana.edu/vol12/issue4/ellison.html
7. Humphreys, L.: Mobile Social Networks and Social Practice: A Case Study of Dodgeball. Journal of Computer-Mediated Communication 13(1), 342–343 (2008)
8. Kaikkonen, A.: Internet on Mobiles: Evolution of Usability and User Experience. Doctoral Dissertation, Helsinki University of Technology, Faculty of Information and Natural Sciences Department of Computer Science and Engineering, p. 48 (2009)
9. Wreblovski, L.: Mobile Web Design Tips: Mobile Should Come First (2010), http://www.techradar.com/news/internet/mobile-web-design-tips-mobile-should-come-first-719677?artc_pg=2
10. Nielsen, J.: Mobile Web 2009 = Desktop Web 1998 (2009), http://www.useit.com/alertbox/mobile-2009.html (retrieved December 24, 2010)
11. Norman, D., Nielsen, J.: Gestural Interfaces: A Step Backwards in Usability in Interactions, September / October 2010, pp. 46–49 (2010)

Context-Aware Places of Interest Recommendations for Mobile Users

Linas Baltrunas, Bernd Ludwig, Stefan Peer, and Francesco Ricci

Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100, Bolzano, Italy
{lbaltrunas,bernd.ludwig,fricci}@unibz.it

Abstract. Recommender Systems (RSs) are software tools and techniques providing suggestions for items, such as movies, CDs, or travels, to be of use to a user. In general, a recommendation can be more compelling and useful if the context of the user is known. For instance, in a travel recommender, the season of the travel, or the group composition, or the motivation of the travel are all important contextual factors that, as a traveler normally does, should be taken into account by a system to generate more relevant recommendations. In this paper we show how a context-aware mobile recommender system for places of interest (POIs) selection can generate more effective recommendations than those produced by a non context-aware version, i.e., those normally provided to the city visitors by the local tourist office. Here we mainly focus on the HCI solutions and in particular in the explanation of the recommendations that are perceived by the user as an important element of the graphical interface.

1 Introduction

Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a user [13]. It is a matter of fact that more compelling and useful recommendations can be identified if the context of the user is known [8] [9]. Here we adopt the definition of context provided by [8]: Context is "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves". For instance, in a travel recommender, the season of the travel or the distance to a Place of Interest (POI) are important contextual conditions that should be considered before suggesting (or not) a tourist to visit that POI.

For this reason, context-aware recommender systems (CARSs) are gaining more and more attention [2], and in particular in the tourism domain [1] [16] [7] [3] [6] [12]. But their supposed advantage and the true impact of the context on the user acceptance of the recommendations must be assessed. Moreover, in order to adapt the recommendations to the user's context one must first elicit all the potential contextual factors that may (qualitatively) influence a travel decision. This knowledge can be obtained by referring to the vast consumer behavior literature, especially in the tourism domain [14]. But this list of factors can only be used as a starting point. In a necessary second step the quantitative dependency of the user preferences – his ratings for items – from each single contextual factor must be modeled. This dependency

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 531–540, 2011.
© Springer-Verlag Berlin Heidelberg 2011

model can be built by acquiring, for a user population, explicit evaluations for some of the POIs under several possible values of these contextual conditions. So for instance, one should acquire the evaluations of a museum, or the likelihood to visit such POI, when the users are traveling with or without children. In this example, "traveling with children" expresses a contextual condition that may be relevant for deciding if the visit to the museum is appropriate or not.

Such data is difficult to obtain because it requires a substantial user effort, since the users must provide evaluations (ratings) of items after having experienced them in several different contextual conditions. For instance, referring to the previous example, one should collect evaluations for a museum if visited with or without children. Hence, it is important to design suitable methodologies for simplifying this data acquisition step still offering enough knowledge about the users' preferences to generate effective recommendations. In a previous paper [5] we described an effective methodology for detecting the quantitative dependency of the user preferences from contextual conditions. The proposed methodology relies on a web based survey application where the users are requested to imagine a contextual condition and then to evaluate a POI taking into account the effect of the context on their decisions, i.e., if the context increases or decreases (or does not impact) the likelihood to select such a POI. With this data we have built a predictive linear model that generates for each contextual situation and POI combination a prediction of the relevance of this POI in that condition. This model is the core computational component of ReRex, a mobile iPhone application that enables the user to: a) select the contextual factors that the system must take into account in the generation of the recommendations; and b) browse the context-aware recommendations identified and displayed by the system.

The system was developed by addressing the previously mentioned challenges, in three steps: first the relevance of several important contextual factors were measured; then in-context ratings were acquired for a population of users; and finally a mobile application (ReRex), running on an iPhone, was designed, implemented and tested in a live user experiment. The first two steps of this process are not described here for lack of space; the reader can find a description of the methodology in [4]. Here we focus on the third step and we illustrate the results of a live user experiment where ReRex was compared to a similar system variant that supports an analogous HCI, but does not take into account the user context for the generation and the explanation of the recommendations. In this case the recommendations are sorted by their average rating among a population of visitors, which is the sorting usually adopted by the local tourist office when offering suggestions to the tourists. An important interaction aspect of ReRex is its explanation functionality (see [15] for a survey on the role of explanations in RSs). Using the above-mentioned predictive linear model, ReRex is capable to identify the most influential contextual condition for each particular recommendation request, and can therefore justify its proposed recommendations by mentioning this condition in the recommendation description. So, for instance, ReRex can detect that the visit to a museum is recommended because the day is rainy and visiting a museum is a good activity for a family in that situation. The live user experiment showed that ReRex's context-awareness can largely improve the system effectiveness and that the users prefer the context-aware variant with respect to the non context-aware one.

The rest of the paper is organized as follows. Fist we illustrate the ReRex user interface and in particular its context-aware recommendation functionality (Section 2). Then, in Section 3, we describe the evaluation methodology, based on a within group comparison of ReRex with a variant, obtained by removing its context-awareness capability. We report there the results of this evaluation and, finally in Section 4 we briefly summarize the outcome of this research and we mention some future work.

2 The ReRex User Interface

In a typical interaction with ReRex the user initially establishes the context of the visit. Using the system GUI the user can enable and/or set the values of context variables representing important contextual factors. The user can switch on/off some of these variables, e.g., the "Temperature" or "Weather" (see Figure 1, left). When one of these variables is switched on the recommender system will take into account their current values in the recommendation generation process. We observe that some of these values are automatically obtained from an external weather forecast service, e.g., the temperature or the weather conditions. Whereas for some other variables the user is supposed to enter a value (with an appropriate screen – not shown here for lack of space). For instance, the user can specify the group composition of the current travel (here called "companion"). In both cases the recommendations are adapted accordingly.

The full set of contextual variables considered in ReRex, their values, and whether they are automatically collected or not is provided in the following:

- Distance to POI (automatic): far away, near by;
- Temperature (automatic): hot, warm, cold;
- Weather (automatic): sunny, cloudy, clear sky, rainy, snowing;
- Season (automatic): spring, summer, autumn, winter;
- Companion (manual): alone, friends, family, partner, children;
- Time day (automatic): morning, afternoon, night;
- Weekday (automatic): working day, weekend;
- Crowdedness (manual): crowded, not crowded, empty;
- Familiarity (manual): new to city, returning visitor, citizen of the city;
- Mood (manual): happy, sad, active, lazy;
- Budget (manual): budget traveler, price for quality, high spender;
- Travel length (manual): half day, one day, more than a day;
- Means of transport (manual): car, bicycle, pedestrian, public transport;
- Travel goal (manual): visiting friends, business, religion, health care, social event, education, cultural, scenic/landscape, hedonistic/fun, activity/sport.

After the user has entered the specification of the contextual situation (see Figure 1, left) the system can be requested to provide recommendations. A short number of suggestions, namely six, are the provided (see Figure 1, right). If the user is not happy with these suggestions he can request "more" of them. This is achieved by touching the corresponding button on the top right of the screen. The user is then requested to specify the type of POIs he is looking for (e.g., castle or art museum or else) and finally more POIs of that type are recommended.

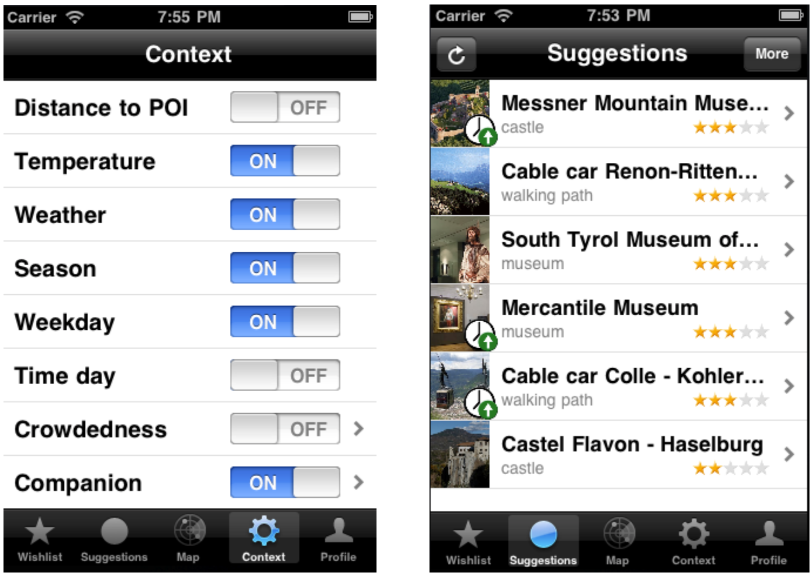


Fig. 1. ReRex screen for context management (left); the user can set (unset) the contextual conditions that the system must take into account in the recommendation generation. The second screen (right) shows a small set of ranked list of POIs recommendations.

In the suggestion list the user can touch any of these suggestions to access a more detailed description of the POI (see Figure 2, left). It is worth noting that some of these suggestions are marked with an icon showing a small clock and a green arrow. This means that these recommendations are particularly suited for the current context of the request as it was previously acquired. In the description of these recommendations (Figure 2 left) there is an explanation sentence like "This place is good to visit even if it is cold today". This refers to the contextual condition that was largely responsible for predicting the top rank for this item. Note, that cold weather condition could even decrease the rank of some items, i.e., their relevance for the current context. However, some items become more attractive than others (indoor museum in our case) if the weather is bad. The other items, i.e., those not marked with the clock icon, are suited as well for the current contextual situation. But we decided not to explain their relationship with the context to highlight and better differentiate those marked with the clock icon from the rest. This can be considered as a persuasive usage of the contextual information.

We have identified custom explanation messages for all the possible 54 contextual conditions listed previously. We note that even if more that one contextual condition holds in the current recommendation session, and all of them are actually used in the computation of the predicted score of each recommendation, nevertheless the system exploits only one of them for the explanation. The contextual condition that is used in the explanation is the most influential one as estimated by the predictive model used by the recommender to predict the relevance (rating) of items in the current context. This design choice is motivated by a simplicity reason; we hypothesized that a single statement would be easily understood by the users and ultimately would produce the best effect on them. Naturally this issue, and more in general a better explanation functionality could be implemented in a future version of the system. In fact, as it will be illustrated in the next section, the quality of these canned explanations are not perceived by the users as strikingly good, indicating that a better explanation message could be generated.

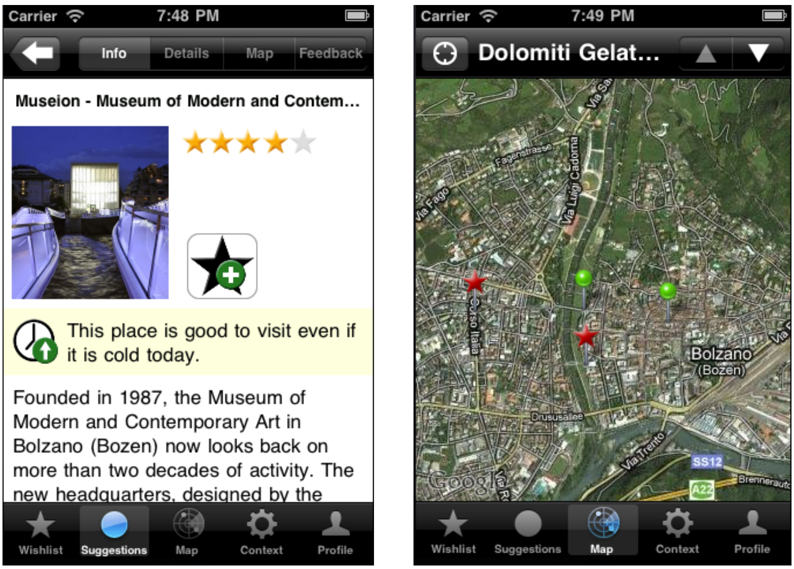


Fig. 2. ReRex screen for explanation visualization (left); the user can browse details of the recommended POI and read the explanation for the recommendation. The second screen (right) shows the position of the recommended POIs on a map.

Some additional functions have been implemented to enable the user to better exploit the system. The first one is called "wishlist" and enables a user to select any recommended POI and add it to a personal wish list. This is activated by touching the star like icon with a green plus shown in the details view of a POI (Figure 2, left). The user can also browse his wish list by touching the "wishlist" button on the bottom left of any screen. The wish list screen is very similar to the recommendations list visualization screen but has an important feature; if the contextual conditions are changing, either because the user is entering new values (e.g., he specifies that he feels "active" now), or because some "automatic" factor is changing (e.g., the weather), ReRex automatically computes a new ranked list and if a POI present in the wish list has increased (decreased) its predicted relevance score the system shows this information with a green upward (red downward) arrow, similarly to those used in the "suggestions" screen (in Figure 1 right). The motivations of these changes are shown in a detailed view of the POI as in Figure 3. This is a useful and compelling characteristic of ReRex that was highly appreciated by the users.

The last function is a typical one for a recommender system, namely entering a rating for a POI. This is available by touching the "feedback" button (top right) that is shown in the POI details screen (Figure 2, left). If the user touches this button a new screen opens letting the user to enter a context-aware rating for the item. This information is sent to the server and will be exploited in the next recommendation requests.

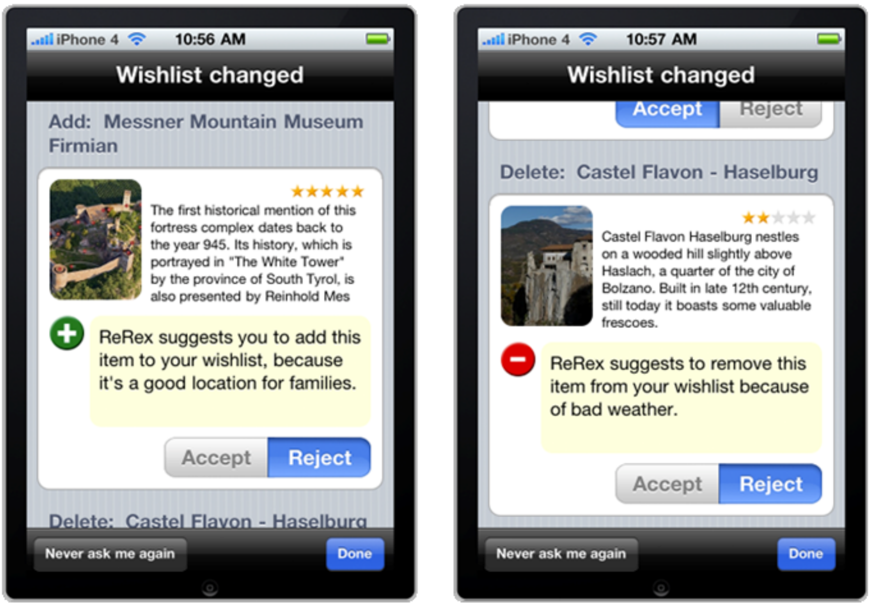


Fig. 3. ReRex screens for visualizing the effect of a contextual situation change in the wish list of the user

We conclude this section by briefly describing the software architecture of ReRex. It is a client/server application where the iPhone client is just managing the human/computer interaction. All the recommendation requests are sent to a server that stores the current user model (the user ratings) and the current request context and computes the recommendations. The client/server protocol is based on a custom designed set of XML files as described in [11]. The recommendations are computed by first computing the predicted relevance score for all the items, depending on the item and the context description. Then the items are sorted according to the predicted score and sent to the client. The complete description of the predictive model is provided in [4].

3 Experimental Evaluation

In order to measure the effectiveness of this approach we developed two variants of our ReRex mobile recommender system. The first is that described previously, the second variant is not context-aware, i.e., there is no possibility for the user to specify the current context, and the UI screen shown in Figure 1, left, has been removed. This variant does not offer any explanation for the recommendation, so the screen shown in Figure 2, left, does not display any explanation message, but it provides the other information about the POI exactly in the same way. The recommendations offered by this second variant are sorted according to the item average rating that we collected previously using a web-based interface (mentioned briefly in the Introduction). Hence, the recommendations produced by this second variant are essentially similar to those that the local Tourist Office makes to a visitor without any special knowledge of the context of the visit.

Using these two variants we could measure if the proposed context-aware system can improve the user perceived relevance of the recommendations with respect to those generated by a non context-aware model. To achieve this goal the test participants, 20 in total, tried out both variants of the system (within groups experimental model), in a random order, and executed, supported by each system, similar but different tasks. For instance, in one of these two tasks the user was instructed to:

> *"Imagine you are living in Bolzano. Suppose that it is a cold, rainy Wednesday and you are alone. Select a single point of interest of your choice and add it to your wish list. Then, suppose that your parents are coming to visit you. If needed, revise your previous selection and add another point of interest to your wish list."*

A second similar, but different task was assigned when the user was asked to try the second variant; hence each user tried both variants and each variant was used to solve one (randomly assigned) of these tasks. After the user completed the assigned task using one system, he was requested to fill out a usability questionnaire including the following statements:

- Q1: It was simple to use this system.
- Q2: The interface of this system is pleasant.
- Q3: The organization of information provided by the system is clear.
- Q4: It was easy to find the information I needed.
- Q5: The system was effective in helping me complete the scenario.
- Q6: It was easy to learn to use this system.
- Q7: Overall, I am satisfied with this system.
- Q8: I like using this system.
- Q9: This system has all the functions and capabilities I expect it to have.
- Q10: I am satisfied with the suggested points of interest.
- Q11: I can effectively find interesting suggestions when using this system.
- Q12: I understood the benefit of using the contextual conditions.
- Q14: I am satisfied with the provided contextual explanations.
- Q13: It was easy to specify the desired contextual conditions.
- Q15: I believe that the contextual explanations are useful.
- Q16: The contextual explanations provided by this system are clear.

The questionnaire was filled out again after the user tried the second system. Statements 12-16 were provided only when the questionnaire was filled out after testing the context-aware version. The user could express a level of agreement to the statement ranging from -2 (strongly disagree) to 2 (strongly agree). These questions were extracted, and slightly adapted to the scope of our investigation, from the IBM Computer System Usability Questionnaire [10]. Moreover, at the end of the interaction with the two variants we asked two additional questions:

- Q 17: which system the user preferred.
- Q 18: which system suggested more appropriate points of interest.

The average response to the first eleven questions, which were asked after the user tested both variants are shown in Figure 4. There is a clear preference of the users for the context-aware variant. However we must note that not all the observed differences are statistically significant (t-test); those significant are 5, 7, 9-11.

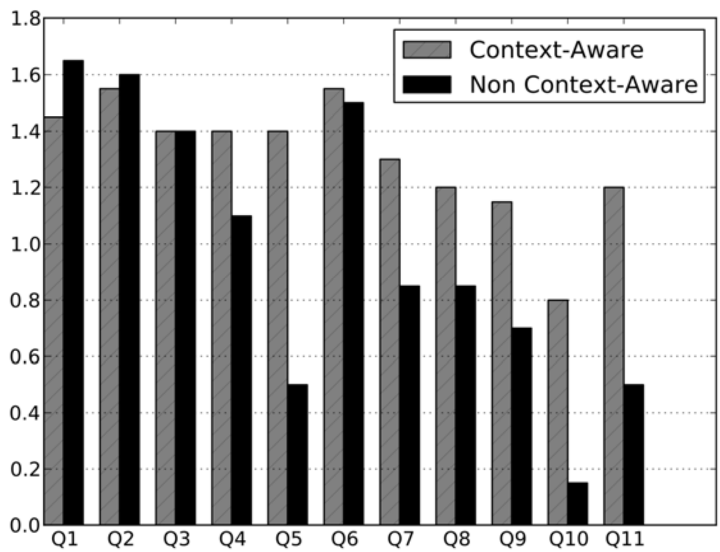


Fig. 4. Average of responses for the first 11 questions of the usability questionnaire. The response values could range from -2 (strongly disagree) to +2 (strongly agree).

We may observe that both systems are considered easy to use and their interface is nicely organized, but the context-aware system is:

- more effective in assisting the user for completing the task (Q5);
- it produces overall higher satisfaction (Q7);
- it is more complete, in term of functionality (Q9);
- it suggests POIs that more largely satisfy the user (Q10);
- it suggests more interesting POIs (Q11).

Regarding questions Q12-Q16, which were asked only after the context-aware variant was tested, the average results are: Q12 1.30; Q13 1.10; Q14 1.05; Q15 1.50; Q16 1.20. So, also with respect to the specific context-aware aspects the users were globally positive. We note that among these aspects the one that it is most interesting is the explanation support. Paradoxically the quality of the explanations scores low (Q14 1.05), but its importance scores higher (Q15 1.50). This indicates that the recommendation explanation is a very important component, and the user is particularly sensible to the quality of these explanations; but the current formulation of the system explanations can be further improved, as these are simple canned texts and no special natural language generation techniques have been used.

Finally, when the users were requested to directly compare the two variants (Q17 and Q18), 85% of the users preferred the context-aware version, while 95% of the users considered the context-aware recommendations more appropriate. In conclusion this evaluation, even if limited in the number of testers, provided a clear evidence that the context-aware recommendations produced by the system were more effective than those produced by the non context-aware version, i.e., the recommendations normally provided to visitors of the city of Bolzano by the tourist office.

4 Conclusions and Future Work

In this paper we have illustrated the importance of exploiting a traveler's contextual conditions when recommending POIs. The proposed mobile application offers to the user context-aware recommendations that are justified and explained by referring explicitly to the contextual situation in which the user will experience them. We have shown that the proposed system can offer effective context-aware explanations that are generated by identifying the contextual conditions that show the largest influence on the predicted relevance score (rating) of the recommended items. In a live user study we have compared a context-aware version and a non context-aware one. We have shown that the user acceptance and satisfaction is larger for the context-aware version and that the users prefer this version compared to another, with a very similar user interface, which does not consider the request context.

As future work we want to address two important issues, personalization and a better explanation functionality. Regarding the first point, we observe that ReRex does not personalize the recommendation, i.e., the recommendations are not based on the domain specific preferences of the user. The system ignores if the user likes more museums of sport activities, and the recommendations do not reflect these preferences. In fact, we do not ask the user to express this type of preferences, and this could also be perceived as an advantage, since it simplifies the human computer interaction. In a future work we want to better understand the role of personalization and contextualization, comparing a context-aware, but not personalized system, as that described in this paper, with a system that generates personalized recommendations, and possibly even combines personalization with contextualization. The second issue was mentioned already in the paper and refers to the measured low user satisfaction for the generated explanations. We want to improve the quality of the explanations exploiting advanced natural language processing techniques to better adapt the explanation to the type of the recommended item and using more information extracted from the predictive model.

References

1. Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R., Pinkerton, M.: Cyberguide: a Mobile Context-Aware Tour Guide. Wireless Networks 3(5), 421–433 (1997)
2. Adomavicius, G., Tuzhilin, A.: Context-Aware Recommender Systems. In: Ricci, F., Rokach, L., Shapira, B., Kantor, V. (eds.) Recommender Systems Handbook, pp. 217–256. Springer, Heidelberg (2011)
3. Ahn, H., Kim, K.-j., Han, I.: Mobile Advertisement Recommender System Using Collaborative Filtering: MAR-CF. Technical Report of The Korea Society of Management Information Systems (2006)
4. Baltrunas, L.: Context-Aware Collaborative Filtering Recommender Systems. PhD Thesis Free University of Bozen-Bolzano (2011)
5. Baltrunas, L., Ricci, F., Ludwig, B.: Context Relevance Assessment for Recommender Systems. In: Proceedings of the 2011 International Conference on Intelligent User Interfaces. ACM Press, New York (2011)
6. Bellotti, V., Begole, J.B., Chi, E.H., Ducheneaut, N., Fang, J., Isaacs, E., King, T.H., Newman, M.W., Partridge, K., Price, B., Rasmussen, P., Roberts, M., Schiano, D.J., Walendowski, A.: Activity-Based Serendipitous Recommendations with the Magitti Mobile Leisure Guide. In: Proceedings of the 2008 Conference on Human Factors in Computing Systems, CHI 2008, pp. 1157–1166. ACM Press, New York (2008)
7. Cena, F., Console, L., Gena, C., Goy, A., Levi, G., Modeo, S., Torre, I.: Integrating Heterogeneous Adaptation Techniques to Build a Flexible and Usable Mobile Tourist Guide. AI Communcation 19(4), 369–384 (2006)
8. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5(1), 4–7 (2001)
9. Dourish, P.: What We Talk about When We Talk about Context. Personal and Ubiquitous Computing 8(1), 19–30 (2004)
10. Lewis, J.R.: IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human Computer Interaction 7(1), 57–78 (1995)
11. Peer, S.: Real-Time Context-Aware Recommendations for Mobile Users. Bachelor of Science in Applied Computer Science Thesis. Free University of Bozen-Bolzano (2010)
12. Ricci, F.: Mobile recommender systems. Journal of Information Technology and Tourism 12(3) (2011) (in press)
13. Ricci, F., Rokach, L., Shapira, B.: Introduction to Recommender Systems Handbook. In: Ricci, F., Rokach, F., Shapira, L., Kantor, B. (eds.) Recommender Systems Handbook, pp. 1–35. Springer, Heidelberg (2011)
14. Swarbrooke, J., Horner, S.: Consumer Behaviour in Tourism, 2nd edn. Butterworth-Heinemann (2006)
15. Tintarev, N., Masthoff, J.: Designing and Evaluating Explanations for Recommender Systems. In: Ricci, F., Rokach, L., Shapira, B., Kantor, V. (eds.) Recommender Systems Handbook, pp. 479–510. Springer, Heidelberg (2011)
16. van Setten, M., Pokraev, S., Koolwaaij, J.: Context-Aware Recommendations in the Mobile Tourist Application COMPASS. In: Adapive Hypermedia 2004, pp. 235–244. Springer, Heidelberg (2004)

Moody Mobile TV: Exploring TV Clips with Personalized Playlists

Arne Berger[1], Robert Knauf[1], Maximilian Eibl[1], and Aaron Marcus[2]

[1] Chair Media Computer Science, Chemnitz University of Technology,
Strasse der Nationen 62, 09107 Chemnitz, Germany
{arne.berger,robert.knauf,eibl}@informatik.tu-chemnitz.de
[2] Aaron Marcus and Associates, Inc. (AM+A)
1196 Euclid Avenue, Suite 1F
Berkeley, California 94708-1640, USA
aaron.marcus@amanda.com

Abstract. An interface for filtering large video repositories for generating personalized playlists via navigation and selection of moods and emotion on a mobile device.

Keywords: mood navigation, customization, interface, query reformulation.

1 Users on the Move

The means of selecting and consuming audio-visual content in mobile settings varies significantly from ordinary domestic usage [1]. Users are in a hurry and have only a limited time available for enjoying entertainment on their mobile smart phones. Whether they are commuting or going out with friends, users usually avoid constructing complex search queries to find suitable content to watch. Besides time constraints, lack of concentration, contextual distractions or physical challenges (e.g., other people watching what they are doing or bright ambient light or too much noise in the subway or having to hold on to a strap in the subway and also to navigate a phone while keeping upright), obtrusive user interfaces, and constrictions from interpersonal circumstances, may prevent a search engine from gaining insight into the entertainment needs of the users.

1.1 Locally Driven Interest and Internet Based Video Consuming Behavior

Studies indicate that interest in local TV content also persists beyond the geographic broadcast limitation of individual stations. “Expats” and commuters eagerly look at “their” local station's website to follow features from their hometown. Not least, this imparts a persistent bond to one's hometown. While the local TV stations have a loyal target audience in large segments of the population, the target audience in the age group “under 30” is increasingly turning to the Internet as the primary medium for their video consumption. We seek to examine how the persistent interest in media coverage from one's place of residence can be combined with the viewing habits of the Internet generation.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 541–548, 2011.
© Springer-Verlag Berlin Heidelberg 2011

We strongly believe that a human component is crucial for recommendation based on emotions. Therefore we created and annotated in a crowd sourcing approach a database of 1000 videos in co-operation with local TV stations in Saxony (Germany) as a basis for personalized video playlists for mobile devices.

2 Simplicity and Plausibility of Video Recommendation

Finding interesting media items quickly and matching the users' entertainment requirements at just the right moment is difficult for both the content provider and the requesting consumers. The lack of both descriptive metadata and broad user feedback adversely affects retrieving suitable content inside a video portal with as few clicks as possible. Most often, the seeking consumers are left having to refine their search query several times or to use non-customized item lists such as "most viewed". Most commonly, retrieving media items referring to a local geographic area, a special interest, or a social group can only be achieved efficiently by following the "channels" of certain users or user groups. Social media sites such as www.youtube.com try to recommend relevant video clips as soon as users finish watching a clip. The system does not communicate the basis upon which a recommendation has been made to the users. Consequently, the algorithm might carry the user farther away from his entertainment source. Furthermore, the scheme on which the recommendation is carried out is a black box for users, leaving them with the experience that successive video clips are way too similar. Variants of the same video clip occur over and over again, or viewers are confused, because consecutive video clips seem to have no connection at all.

On the other hand, traditional TV stations handle the task of connecting single broadcasts very well. A moderator weaves a golden thread and guides viewers via this potentially emotional connection through a series of video clips. This TV format is well established and favorably to the algorithmic approach described above. It is, however, not readily transferable to the generation and consecutive play-out of personalized playlists.

Instead of providing a personal moderator, we introduce a semantic framework that enables the viewer to weave his own golden thread to select and connect interesting items. By applying readily available metadata from our database, this may be approached by selecting a list of video clips via time, genre and broadcasting station. This would allow answering four possible content searches: **When** did it occur, **What** is it about, **Where** did it happen and **Who** recorded it. However, all four lack any of the moderator's emotional features. For example, video clips about a traffic accident and offspring at the local zoo may follow right after each other in a playlist, just because both happen to occupy the same genre and happened around the same time. Since a human's moderation is lacking, the sequence might make sense, but may be disturbing for the audience.

3 Selecting Content vs. Constructing Playlists

Approaches to customize a content stream depending on users' interests usually learn by analyzing content description, genre data and users' viewing history. Consequently, they enable entertainment systems to adapt themselves to their owners' long-term habits. What they do not consider is the fact that peoples' individual moods vary at frequent and irregular intervals. To meet these requirements, media items need to be classified using emotional descriptions along with giving viewers an instrument to express their current mood.

Well-established systems that allow users to select content based on mood or emotion concentrate greatly on a single domain. To a certain extent, they are capable of plausibly recommending thematically related media objects as a playlist. Services such as www.last.fm, www.putpat.tv, or www.tape.tv merely contain music videos. Regarding this specific genre, recommendation is simply achievable, because (due to the domain limitation) it is based on a similarly limited plausible vocabulary. On the site www.tape.tv one can select material according to the genre or a mood. However, the moods cannot be chosen from a complete list of all imaginable possibilities, but from a list generated by the site's editors. It is also peculiar that one cannot combine several of these filter parameters. On the other hand, the site www.putpat.tv uses several slide controls (the parameters of which can be randomly occupied) to "adjust the station". These degrees of freedom enable competing queries such as "more Madonna, less pop"; at the same time. The concept of self-configurable sliders encourages users to try different things but, again, lacks an extensive list of moods to choose from.

While www.putpat.tv and www.tape.tv introduce novel, yet rudimentary, emotion based content filters, navigation functionality remains limited, compared to those available on www.youtube.com or www.last.fm. The first-mentioned two allow to skip the video clip currently playing in order to "ban it forever" or "love it", while the latter two evaluate user interaction and choices to aid the underlying algorithm in its learning. Both concepts are solely based on recommendation strategies based on an initial query in connection with users' listening histories.

All introduced services lack a well-defined emotional vocabulary and focus on mere algorithm based decisions. Both restrictions render the concept of a personalized playlist highly rigid and not customizable. In order to accumulate content based on emotion or moods, two main aids are missing in the concepts described above: First, the aforementioned list of emotions and second, a set of navigation options to allow a real personalization.

4 Using TV-Anytime to Contribute Emotional Metadata

In the recent past services for the purpose of TV personalization has been proposed. Obviously, all of them are subjected to descriptive and structured metadata. To unify this common feature, the TV-Anytime standard has been introduced to define metadata schemes, especially concerning genre and broadcast related data. [2] Subsequently, the standard influenced research efforts on TV services driven by users' needs, such as [3.]. Generally, emotional classification has been left out as the basis of recommendations is still limited to genre data, content descriptions and users' long-term viewing history. Moreover, the TV-Anytime standard offers an emotion-based content description scheme consisting of 53 adjectives, which has been proved in reflecting emotional impacts to the viewers [4., 5.]. Regarding the advantages of a cross-system standard describing emotions and providing an anticipated well-structured order, our work takes a closer look at the suitability of the given listing.

5 Implementation and Evaluation

5.1 Design Methods

The creation of a personalized playlist is "more of an Art than a Science" [6]. The mere application of an algorithm is not enough for most users. [6] Therefore we adapted and applied a couple of design methods in order to gain insight into potential users mindsets regarding the construction of personalized playlists and the use of content filters, especially mood filters in a TV context. First, we asked participants to draw cognitive maps of possible ways for filtering video repositories down to satisfying playlists. This was done to get an insight how users think and associate when formulating content filters. Second, based on the subsequent results we crafted paper prototypes and discussed the sketched out possibilites for navigating and selecting with potential users in open-ended questionaries.

Questions we wanted to address are:

Which means are convenient for users to individually select video clips as a playlist?
Which means are necessary for users to perceive these playlists as such?
Is mood a key mean for that?
Are the moods suitably proposed by the
TV-Anytime subset for the local TV domain?

5.2 Moody TV - Navigation

With moody TV users shall be able to select and combine content filters fluently and independently. To find out what content attributes users are looking for, we asked participants to map out a virtual space of content properties and show how they thought to navigate within it. This method usually helps to discover pathways and interests in which people make sense of a particular content space. The results eventually help to make sense of how to construct queries for filter specification. Users were asked to individually draw a map or diagram of what comes to their mind when being on the move and having a mobile video handset available, whether sitting on public transportation alone or being in a pub with friends. The six users, four male, two female, between the age of 23 and 34 are all regular viewers of local TV. They had 15 minutes time to draw a map or scheme and were asked to freely associate parameters to form a personalized playlist. A discussion with all participants followed.

The results lead to the assumption that users indeed are interested in direct emotional filters. Most of the user generated maps feature emotion clusters or the simple question “how” in a list of questions. See Fig. 1 and 2 for excerpts of users cognitive maps.

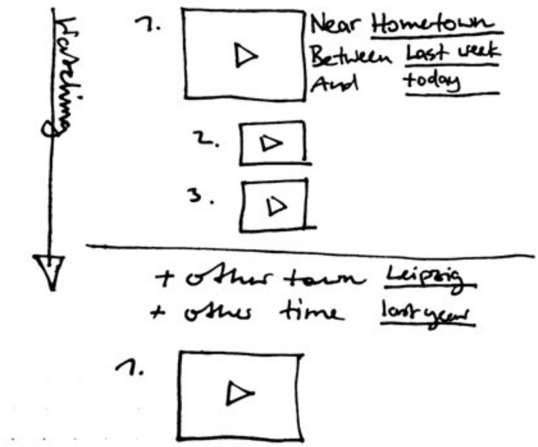


Fig. 1. Cognitive map drawn by a user

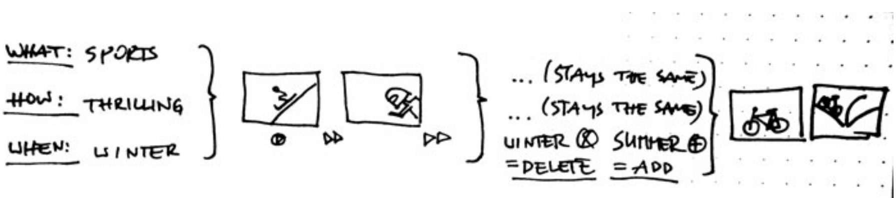


Fig. 2. Cognitive map drawn by a user

Given the mindset of being on the move, users formed questions from a simple vocabulary and subsequently wanted to change only certain parameters after watching a few video items. Although five users out of six proposed emotion in some way as a mean to filter, the results are far from being readily transformable into an interface. It is plausible to assume, that emotion is a relevant feature. However, cooperating with linguists may be necessary in the future to optimize semantic significance of the description vocabulary, which is for now beyond the scope of this work.

5.3 Moody TV - Prototyping

Based on the findings about emotion filters described above, we proposed a low fidelity prototype containing a filter named “How” together with more filters based on the other four cardinal questions Who, Where, When, What mentioned above.

To prove the concept we introduced it to twelve users between the age of 17 and 44. We rapidly sketched a series of mobile device screens with the main objective of constructing a personalized playlist with five clicks at the maximum. See Fig. 3. Findings are discussed in depth in [7].

Fig. 3. From [7]

Users' feedback on this approach was insightful in two ways. On one hand, users at large expressed their general approval on the advantages that might arise by constructing exhaustive content filters with just a few steps of interaction. On the other hand, the pre-structured characteristic was heavily criticized. The rigidly defined prototype inspired participants to incredibly rich feedback. It may be generalized as following:

All twelve users criticized the seven proposed moods as too general. Most users proposed mood sub-categories. *The landscape of moods in users' heads needs to be supported via a detailed browsing option for moods.*

Subsequent questions revealed that users have significantly varying cognitive styles. Although not statistically significant, at least two different styles have been discovered. *The decision to categorize moods is subject to the inclusion of different cognitive styles.*

Applying plain cardinal questions as categories has obscured the users. Six of our test subjects asked for a brief explanation of the questions and their combination. *We conclude that an interface that asks questions, should ask in a more conversational way.*

Three users criticized the interface proposal as too rigid for supporting granularity. Most users discovered that some categories need sub-categories, while others did not agree to that. *The interface needs to support this.*

Three users missed an option to sort the subsequent playlist. This could be addressed via the order in which users answer relevant questions.

The proposal of a low fidelity prototype in combination with open-ended questions has proved to be a fast and convenient way to gain user feedback on a large variety of issues without a lot of explanation. The main insight is, that all users found and used the filter option “how”. Most user feedback was given on only this feature.

6 Future Work: Moody TV - Emotion

Contrary to the proposed prototype a complete emotional vocabulary is featured in [2]. While a set of 53 emotions is sufficient to describe every kind of audiovisual data, it has not been widely implemented for selecting emotions, due to the lack of predefined sets of metadata. When content is tagged based on those metadata, the mere number of possible emotions makes it hard to subsequently navigate and select them in the form of, e.g. a list, a menu or even a tag-cloud.

Up to this point, the system has been converged into an interface that asks questions in an order that is determined by the user.

Based the previous result that users’ cognitive styles are different, the next step is to adapt a design method to sort these emotions into categories that make sense to users in order to navigate the 53 option with ease.

Acknowledgement. This work was accomplished within the project sachsMedia – Cooperative Producing, Storage and Retrieval, funded by the program Entrepreneurial Regions of the Federal Ministry of Education and Research, Germany (FKZ: 03IP608). The authors are responsible for the content of this publication.

References

[1] Schildbach, B., Rukzio, E.: Investigating selection and reading performance on a mobile phone while walking. In: Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, pp. 93–102 (2010)

[2] TV Anytime, pp. 1–160 (December 2010), `http://www.etsi.org/deliver/etsits/102800102899/1028220301/01.02.0160/ts1028220301v010201.pdf`

[3] Zhang, H., Zheng, S.: Personalized tv program recommendation based on tv-anytime metadata. In: Proceedings of the Ninth International Symposium on Consumer Electronics (ISCE 2005), pp. 242–246 (2005)

[4] Lee, H.K., Kim, J.G., Yang, S.J., Hong, J.: Personalized tv services based on tv-anytime for personal digital recorder. IEEE Transactions on Consumer Electronics 51(3), 885–892 (2005)

[5] Petersen, M., Butkus, A.: Extracting Moods from Songs and BBC Programs Based on Emotional Context. International Journal of Digital Multimedia Broadcasting (2008)
[6] Cunningham, S.J., Bainbridge, D., Falconer, A.: More of an art than a science: Supporting the creation of playlists and mixes. In: Proceedings of the 7th International Conference on Music Information Retrieval, pp. 240–245 (2006)
[7] Berger, A., Knauf, R., Eibl, M.: Constraints and simplification for a better mobile video annotation and content customization process. In: Workshop Proceedings of the EuroITV, 1–4 (April 2010)

Mobile Web and Native Apps: How One Team Found a Happy Medium

Kathleen Buettner and Anna M. Simmons

Design Innovation Group, Intuit Inc., USA
{kathleen_buettner,anna_simmons}@intuit.com

Abstract. Mobile web and native apps bring different pros and cons to the table. Intuit recently embarked on a journey to create a mobile web version of its online business banking solution. In developing this product, the team attempted to get the best of both worlds by creating a system that would recognize a device class and adapt content to better fit it, providing a more consistent user experience while still taking advantage of the wide accessibility gained by utilizing the mobile web.

Keywords: Mobile, iPhone, Android, Mobile Web, Blackberry, HTML5.

1 Introduction

The number of web-enabled mobile devices is growing at an astounding rate. By early 2010, the number of smart phones worldwide had already exceeded 4.6 billion. Morgan Stanley estimates that by 2012 smart phone shipments will surpass that of personal computers [3]. As a result, developers are working rapidly to create products and services optimized for this newly prevalent platform. Currently, there are two main choices when it comes to delivering your content on a mobile device: a native app or a mobile website. (Historically, hardware manufacturers provided another way to get content onto mobile devices, but this model is on the decline as smart phones continue to flourish.)

A native application (or app) is an application that was created for a specific device. It can take advantage of device capabilities (such as the camera, GPS, etc), adhere to device-specific design conventions, and is typically downloaded from an app store or marketplace managed by the OS provider. A mobile website, however, is just a website that (ideally) has been optimized to fit nicely on a mobile screen, function efficiently with cellular network speeds, and accommodate mobile navigation controls (such as a finger or a trackball).

Not surprisingly, both of these categories are growing at rapid rates. According to a report by Taptu, the number of websites optimized for mobile touch devices specifically (i.e. those smartphones that utilize a touch screen vs a trackball or keypad for primary control) grew 35% to reach 440,100 sites between the months of December and April in 2010, suggesting a projected annual growth rate of 232% [2]. Compare this to Apple's App Store, which hosted 185,000 apps and is growing at an annual growth rate of 144% [2]. The newest addition to the party, the Android

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 549–554, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Marketplace, offered 35,947 apps in April, leading to a projected annual growth rate of a whopping 403% as it starts to get off the ground [2]. While the Apple App Store and Android Marketplace both continue to grow at impressive rates, the mobile web contains more sites than both app stores combined.

So when is it appropriate to use a native app and when should one utilize mobile web? What are the differences? As one might expect, there are pros and cons to each delivery method.

1.1 Advantages and Disadvantages of Native Apps

Native apps have a lot to offer content developers. In terms of pure performance, native apps lead the way in speed and efficiency for the user. Native apps also give the user the option of accessing content without necessarily being connected to a cellular or wifi network. Since apps are attained via an app store or marketplace, there is a clear process and expectation set already regarding findability and exposure. The app store has a search mechanism in place, as well as “top app” lists that help users find offerings. Additionally, certain offerings may be “featured,” giving them more prevalence in the store. Similarly, a monetization model has already been created and the payment mechanism is already in place. In most cases, the user purchases the app through the app store itself, so the product teams need only worry about setting a price. Furthermore, a native app allows you to take advantage of device attributes like the camera, photo albums, and GPS feature, to name a few, which can add increased functionality to your offering. For all these reasons, the native app route can be very compelling.

That said, there are also some disadvantages to creating a native app. While the app store can be great for findability and promotion of an offering, it also puts apps side by side with competitors and opens them up to the possibility of negative reviews. Furthermore, an app store or marketplace does not enable the same amount of reach as a mobile website, since one must build a different version of the app for each OS. Also, with an app store comes app store regulations. This can affect cost (developer fees and profit), scheduling, and content, since that third party must approve the offering. For example, approval by Apple’s app store might take two weeks or more. This means that even though developers may have finished coding an update, they possess fairly limited control over when the users will actually get to see it. Finally, the Apple app store will take a cut off the top of each app sale.

So when might one choose to create a native app? Native apps work best when an offering requires complex interactions, lots of processing power, and access to device hardware (such as the camera). Therefore mobile games, like Rovio’s Angry Birds, lend themselves well to being native apps; achieving the necessary performance required to support the optimal gaming experience would just be difficult to do in a mobile web environment. Likewise, an offering such as Intuit’s SnapTax that utilizes the device’s camera as an information entry method makes sense as an app because access to that hardware is not yet possible from the mobile web platform.

1.2 Advantages and Disadvantages of Mobile Web

Creating a mobile website has its own set of advantages and disadvantages. Utilizing a mobile website allows the offering to get universal exposure – it can be accessed from any web-enabled mobile device and take advantage of knowledge of a branded website the offering may already have. Additionally, with a mobile website, development releases can happen on your own timeline, and instantaneously, versus waiting on an approval and release cycle like an app store might require.

While mobile web arguably makes it easier to get widespread access from multiple platforms, it cannot function offline and access device attributes in the same way that an app can. Plus, since the website will be accessed across platforms, it cannot take advantage of design conventions for a specific device in the same way that an app can. For example, offerings designed specifically for an iPhone can look quite different from offerings designed specifically for an Android device; each operating system follows its own design rules and utilizes elements of the hardware (such as buttons) in a way that the user comes to expect. When an offering is developed for the mobile web, its design has to work on a broad range of devices, and hence cannot cater to any one specific set of design assumptions.

One of the biggest representatives in the mobile web space is retailers [2]. Large stores want to capitalize on market share across multiple device platforms; they need to reach the maximum amount of customers with the smallest development effort. Being on the web also allows them quick and easy access from a web search, something the user would likely be doing anyway in a shopping task. Combine these factors with the ability to instantly update the mobile web product without third party approval and profit sharing, and it becomes clear why the mobile web is a tempting option for retailers (and others!).

2 Intuit's Mobile Web Case Study

A team at Intuit was tasked with coming up with an elegant mobile offering for Intuit Financial Services' online business banking product. Though the offering was already a web product, the core website did not display in a usable manner on a mobile device. This was not a simple task given that the online solution was complex and would need substantial work in order to package and deliver on mobile.

To determine whether to go the mobile web or native app route, the team considered the mobile devices already being used by their current customers. They found that there was no single dominant device to target. This meant that, were the team to choose a native app, they would need to develop an app for each of the predominant mobile app platforms in order to serve their customers. This would involve considerably more engineering effort than building one mobile website. To allow them to reach current customers while also minimizing engineering effort, the team decided to build a mobile web offering. The team understood that this meant giving up a wholly consistent user experience, since the offering could not be developed to take advantage of device specific design conventions.

In an effort to compensate for variations across the multitude of devices, the team opted for a device targeted content adaptation model. To achieve this, the team looked across the breadth of phones on the market and reviewed their characteristics. This enabled the team to break the phones into categories or “classes” based on common characteristics most relevant to the product: similar screen size and ratio, similar method of input of information (e.g. QWERTY keyboard, touch screen, trackball), and similar browser/OS capabilities to support markup and scripting. The first version of the site was developed for two classes. Class 1 included touch screen devices like Androids and iPhones; Class 2 included primarily RIM (Blackberry) devices. They then created a specialized version of the front end website for each device class that was optimized for the characteristics present in those devices. For example, the Class 1 site featured larger target areas so that a finger could easily select them while the Class 2 site featured smaller target areas that could be quickly navigated via a trackball.

Fig. 1. Content Adaptation for an iPhone and a BlackBerry

This development strategy allowed the team to create a user experience that would feel more consistent for users with a touch screen device as well as users that have a trackball device, but did require the website to display differently based on the type of device accessing it. To make this work, the team had to come up with a way to identify which type of device was viewing the site so the appropriate version of the website could be presented. This was solved by a device recognition service that looks at specific device and service configurations as soon as the user requests the first page of the experience. Using this information, the system can deliver HTML, graphic assets, and script that will work well in the device requesting it.

This meant that the site had to be designed in such a way that the appropriate design components could be pulled in once the device class was identified. Rather than creating new graphic elements for each page of the experience, the team employed a modular design for the site. Any page layout is a compilation of modular components that stack together in a HTML page. Breaking the layout into smaller pieces (modules) makes the adaptation more manageable and easier to enhance in the future. For example, the section header component (see Figure 2) can be utilized throughout the site and in a variety of ways. Also, having reusable modules reduces the amount of CSS and script needed to load in any given page when compared to a custom screen based design, and load time is especially critical in the fast-paced world of mobile.

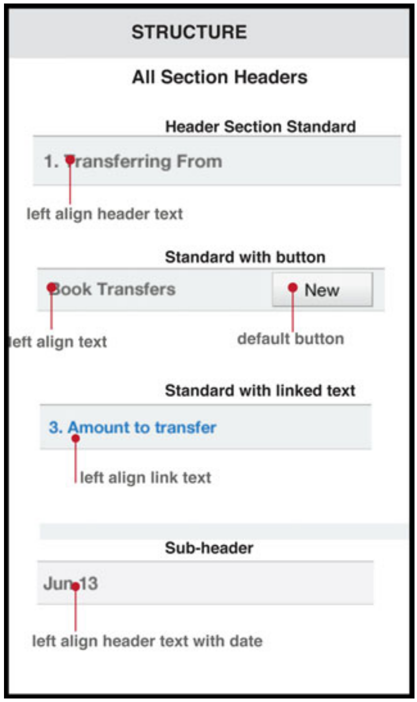


Fig. 2. Modular Design Components

In summary, choosing mobile web allowed the team to maintain a wide reach with their customer base and minimize engineering efforts. By creating different design components for different device classes and a way for the server to identify which class of device is accessing it and hence plug in the appropriate modular assets, the team was able to display a fast-loading, semi-optimized version of the page for users on different devices, helping to make up for some of the typical shortfalls of the mobile web.

3 The Future of Mobile Web and Native Apps

Choosing a mobile strategy is no doubt a difficult endeavor, especially as the mobile marketplace continues to welcome new hardware and software in the form of not only new phones but also tablets, e-readers, and the like. It has been suggested that a browser-based approach is key for reaching the widest number of users (and indeed, that is the reason Intuit chose this platform for their mobile banking product) but, given all the other advantages and disadvantages to each, the battle between mobile web and native apps is far from over. In the web realm, HTML5 promises some critical integrations with mobile device features that could help mobile web take advantage of some of the luxuries of the app world. Still, arguably, the ideal would be a hybrid experience that utilizes the best of both web and app worlds. Many native apps, such as Twitter and Facebook, access the mobile web for at least some content, giving them the ability to access on-device capabilities and UI conventions as well as take advantage of the app store but also access their web services and provide the ability to instantly update content. As advances continue to be made in the mobile web space, this option will become even more powerful. Teams should continue to think about how they can utilize the positive elements of both native apps and mobile web in their future mobile endeavors.

References

1. Adobe Mobile Experience Survey: What Users Want from Media, Finance, Travel & Shopping. Adobe Online Marketing Suite and Adobe Scene7 (October 2010)
2. "The State of the Mobile Touch Web" Taptu (April 2010)
3. "Internet Trends" Morgan Stanley. CM Summit – New York City, June 7 (2010)

Investigating the Integration of Hand-Held Haptic Devices in Daily Work Activities: The Case of a Tennis Coaching Assistant on iPhone

Evanthia Faliagka[2,4], Petros Karkoulias[1,3], Maria Rigkou[1,3,4], Spiros Sirmakessis[1,2,3], Giannis Tzimas[2,4], and Athanasios Tsakalidis[4]

[1] Hellenic Open University, Parodos Aristotelous 18, 26335 Patras, Greece
[2] Department of Applied Informatics in Management & Finance, Technological Institution of Messologhi, Nea Ktiria, 30200 Messolonghi, Greece
[3] Research Academic Computer Technology Institute, N. Kazantzaki str., Rio Campus, 26500 Patras, Greece
[4] Department of Computer Engineering and Informatics, University of Patras, Rio Campus, 26500 Patras, Greece

faliagka@ceid.upatras, pkarkoulias@gmail.com, rigou@ceid.upatras.gr, syrma@teimes.gr, tzimas@teimes.gr, tsak@ceid.upatras.gr

Abstract. This paper investigates the integration of hand-held haptic devices in daily work activities in the form of personal digital assistants (PDAs) and smart phones that help professionals perform their tasks in a more effective way. Such devices provide significant advantages of high portability, availability and familiarity, since their users already use them as mobile phones. The aforementioned issues will be studied in the context of a Tennis Coaching Assistant, an application running on an iPhone that provides management and scheduling capabilities for tennis trainers and allows a comprehensive overview of the execution of on-court drills and coaching programs. The paper describes the usability evaluation methodology applied and documents the results focusing on interaction difficulties and practical obstacles reported by the users as well as suggestions for future versions of the application, which generally received quite positive feedback.

Keywords: iPhone, Haptic, Usability, Evaluation, Objective-C.

1 Introduction

Nowadays we are witnessing a burst of new information mediums and communication devices that users can carry on them and use for various aspects of their daily activities. The interaction styles supported by such devices are not restricted to physical button activation typically used on mobile phones. They also include direct manipulation and more recently, haptic interaction via touch screens, which offer an intuitive way of interacting but at the same time, a rather ‘slippery’ one, especially when dealing with small-sized displays.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 555–563, 2011.
© Springer-Verlag Berlin Heidelberg 2011

This paper investigates the integration of hand-held haptic devices in daily work activities in the form of personal digital assistants (PDAs) and smart phones that help professionals perform their tasks in a more effective way. Such devices provide significant advantages of high portability, availability and familiarity, since their users already use them as mobile phones. Although it is logical assumption that keeping the already available style of operating such devices and following the interaction conventions of basic phone functions provides an initial safe basis for designing the GUI and the interaction style of more special-purpose applications, more thorough study is also required. After all, given that a certain device is effective for a set of trivial tasks it is not necessarily true that the device supports equally efficiently other tasks that oppose different and perhaps more complex operations. Furthermore, it needs to be verified that integrating such an application in the daily work habits of professionals will not impose on them extra effort and delays due to handling difficulties and usability pitfalls.

The aforementioned issues will be studied in the context of the Tennis Coaching Assistant (TCA), an application running on a typical iPhone that provides management and scheduling capabilities for tennis trainers and allows a comprehensive overview of the execution of on-court drills and coaching programs. Coaches can use the application to set up practice programs of varied duration and assemble an overall scheduling by mapping each program on a set of time slots at user defined dates and associating them with a certain athlete or group of athletes. The design of the application GUI and all interaction conventions followed the official Apple guidelines [1, 2, 3, 4] and iPhone UI elements and was evaluated against iPhone interface usability problems identified in the related scientific literature [5, 6, 7].

The rest of the paper is structured as follows: section 2 provides a brief overview of the TCA functionality, section 3 describes the applied usability evaluation methodology (i.e. the composition of the user sample, the questionnaire and the experimental use of the system based on seven specific scenarios, as well as details on the field observation conducted). Section 4 presents the analysis and concluding remarks of the study, while section 5 summarizes the paper.

2 The Tennis Coaching Application

The implemented application can be used by professional tennis trainers to set up practice programs for certain athletes or groups of athletes. Each program comprises a set of practice sessions, each of which (i.e. practice session) is assigned to a specific date and time. A practice session is composed of a series of consecutive drills characterized by their duration, type of skill improvement, level of difficulty, sketch/photo, as well execution instructions. The coach can inspect on his/her iPhone screen all available drills at the time of the practice, along with their attributes. Our application provides management and scheduling capabilities, allowing the trainer to create and modify training programs, as well as assemble an overall scheduling by mapping programs on time slots and associating them with athletes.

The application's GUI is depicted in Figures 1-4, as a set of screens that implement representative functionality. The three main functions (namely calendar, programs management and drills management) are pinned on the horizontal bar for easy access. All GUI elements and options follow the design conventions and the interface design of the embedded iPhone applications. In **Fig. 1** and **Fig. 2** we can see the scheduling function, where the user can see the practice sessions for each day projected on a typical iPhone calendar view. Selecting a practice session (**Fig. 2**), the coach can see the drills comprising it and is able to remove/add drills or change their sequence and duration. He/she can also create a new practice session using a stored template or generate a custom session from scratch. In **Fig. 3** we can see the screen of a specific drill presenting information such as the drill name, its type, its purpose (in terms of skill acquisition), a short description of execution and an image showing graphically the steps of the drill.

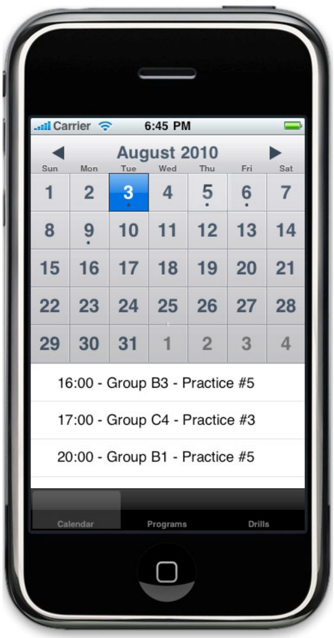


Fig. 1. The calendar application screen

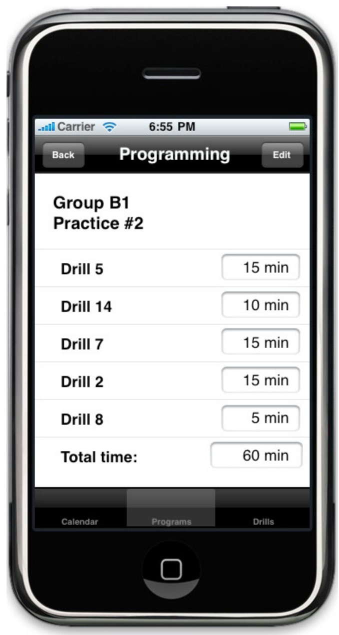


Fig. 2. The program details screen

Fig. 4 presents the screen with all scheduled practice sessions in the selected program (i.e. program GroupA3, comprising tennis athletes of intermediate skill level scheduled to have a series of 15 practices). A program is characterized by its name (which usually corresponds to the name of the athletes' group), its difficulty level (which also refers to the level of tennis skills of the group athletes) and its duration (in terms of how many practices). Selecting a practice on the screen Program Details screen the user can see which drills it contains with their scheduled durations and edit them. The user can also add a new practice or delete an existing one.

At a technical level, the TCA is an object-oriented application developed in Objective-C, the programming language of iOS, Apple's mobile operating system used by iPhone [8, 9, 10]. The UI was developed using the Cocoa Touch collection of frameworks which provides the tools required to implement graphical, event-driven applications in iOS, as well as special GUI controls, buttons, full-screen views and control of the multi-touch gesture interface [11].

Fig. 3. The drill details screen

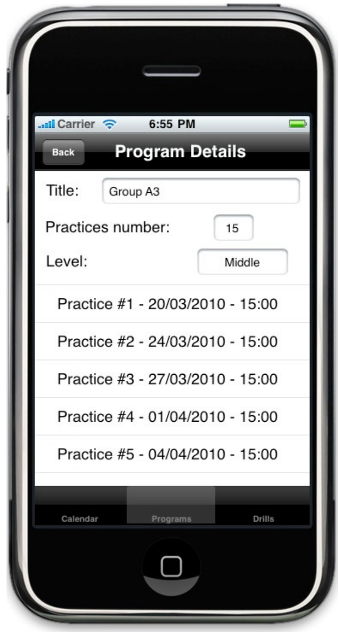


Fig. 4. The program details screen

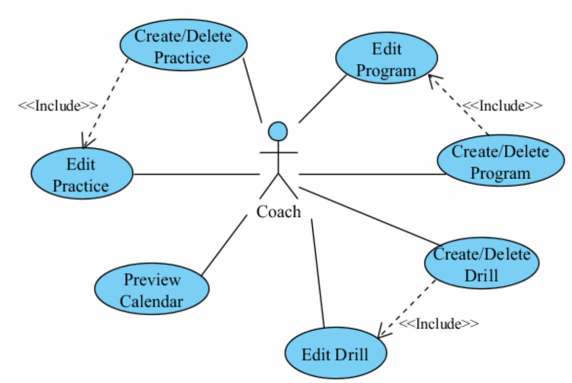


Fig. 5. UML use case diagram depicting all 7 use cases comprising system functionality

The analysis and design that drove the application implementation followed the ICONIX methodology, a software development process lying between the very large Rational Unified Process (RUP) and the very small eXtreme programming approach (XP). The ICONIX process is use case driven and relatively small but it doesn't discard analysis and design like XP does. Moreover, it makes streamlined use of the Unified Modeling Language (UML), while keeping a sharp focus on the traceability of user requirements [12]. This ensures that the development process prioritizes a user-centered approach by expressing the functional requirements of the user in terms of a set of 7 use cases that the final application will be tested against during system evaluation and are presented in **Fig. 5** in the form of a UML use case diagram. Note that even though there is not a one-to-one mapping between use case and test scenarios, test scenarios generally combine 2 or more use cases which increases their difficulty but also provides more reliable feedback on user clarity, ease of use and ease of learning.

3 Usability Evaluation Methodology

3.1 User Sample

The application was evaluated by a user sample consisting of 30 professional tennis coaches. All the coaches selected were representative of end users of the system under study. The user sample was also selected based on the level of their familiarity with the iPhone device and applications, so as to investigate the usability of our system in terms of different types of users. The users that rated our system were 25 men and 5 women, at an age range of 25-53. The user group was divided into two groups based on their familiarity with the iPhone device. The first group consisted of 15 users who were not familiar with the iPhone device and its applications, while the second group consisted of iPhone users (all users of the second group owned an iPhone and had been using it as a phone at least for one year prior to the test). In an attempt to also investigate the usability of the application independently of the handheld haptic device, half of the users (from both groups) were asked to evaluate the application using a simulation that projected an iPhone device on a laptop screen and supported full functionality via a mouse while the others tested the application on iPhone devices. All users were asked to use the application during their daily practice sessions for a week.

3.2 Scenarios

All users were asked to complete the following scenarios (i.e. selected courses of action in the respective use cases of **Fig. 5**):

Table 1. Scenarios for user testing

Scenario 1	Create a new drill (type in all details and add an existing image to it)
Scenario 2	Change the duration, the image and the description of an existing drill (the new image should be imported as a photograph taken using the iPhone at the time of the test)
Scenario 3	Create a new practice session comprising 4 drills
Scenario 4	Create a new program, add 2 practice sessions to it and assign them to specific dates and time
Scenario 5	Open the application calendar, preview a certain day's practices, select a practice and delete it
Scenario 6	Locate the 4th practice of a specific program and shift its date 2 days later and its starting time 3 hours earlier.
Scenario 7	Open the application calendar, preview a certain day's practices, select a practice, preview its drills and replace the 2nd drill with the 4th drill and vice versa.

3.3 Questionnaire

After an initial small introduction to the functionality users were asked to complete the above usage scenarios. The users were also asked to use the application during their daily practice sessions for a week and then to fill in a SUS questionnaire [13] developed to determine the system's perceived usability. The System Usability Scale (SUS) is one of the most popular standardized usability questionnaires typically comprising 10 items, each one of them being a statement (positive or negative) and asks for a rating on a five-point scale from "Strongly Disagree" to "Strongly Agree". In our case it was adapted by replacing the word "system" with the word "application". Using this questionnaire, we tried to determine the application's effectiveness (i.e. whether the objectives were achieved), its efficiency (the effort needed to achieve those objectives) and the user satisfaction.

Moreover, the SUS questionnaire was expanded with some open-ended questions. These questions helped us determine whether the users experienced problems with the lighting or other conditions and get a general evaluation of the application's usefulness and practical deployment. We also examined how well our application integrates with a line of work and how much it increases their productivity. The results are presented in section 4.

3.4 Field Observation

In addition to questionnaires, 7 coaches participated in field observation sessions and used the application during tennis practice, on court. These coaches belonged to the second user group (the users that already owned an iPhone 3G). They were asked to use the application as much as possible in a week's period but at least for two practice sessions each and then were interviewed and provided a set of drawbacks/limitations, as well as suggestions for improvements. The findings of the interviews are presented in section 4.

4 Evaluation Results

The most important metric to quantify the usability of our application was the number of scenarios completed by the test users. The first group (the ones that hadn't used an iPhone device or an iPhone application before) completed 5.9 scenarios on average, while the second group completed 6.6 scenarios on average. Overall, as we can see in **Fig. 6**, the majority of the users (47%) completed all seven scenarios. Also, a high percentage (37%) managed to complete 6 out of 7 scenarios an encouraging observation as users completed almost every task they were assigned.

Fig. 7 depicts user responses to the 10 questions of the SUS questionnaire. All questions were scalar with 5 possible answers each, ranging from "totally disagree" (0) to "totally agree" (4). We can see that users belonging to the second group responded slightly more positively than users belonging to the first group. The second group found the application easy to learn, didn't need support for their fist usage and felt very confident while they were trying to complete the scenarios. Both groups could use the application on a frequent basis and none of them found it complex or inconsistent. The only thing that should be taken into consideration is that the users of the first group needed to learn the basic functionality before the first use and needed a little support during the pilot scenarios as they were not accustomed to the interaction conventions of iPhone.

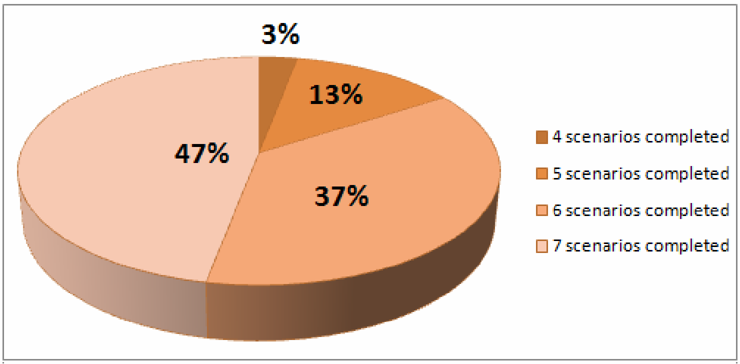


Fig. 6. The scenarios completed

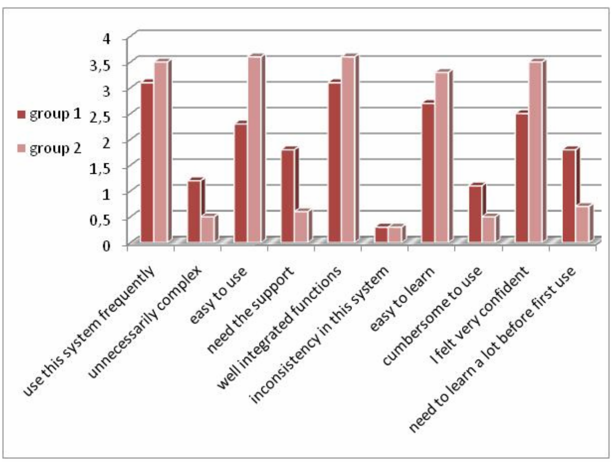


Fig. 7. SUS questionnaire responses

After a week's usage period we investigated how well the application integrated with the work tasks of a coach. We asked the tennis trainers in how many practice sessions they used our application. The answers showed that those who owned an iPhone device used the application more frequently than the trainers who used the application on a laptop computer. The results are summarized in **Fig. 8**

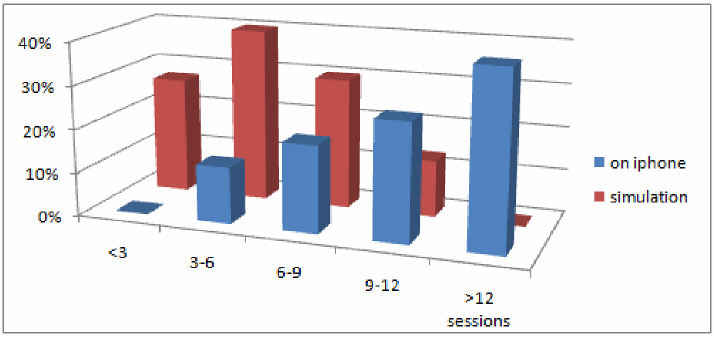


Fig. 8. Frequency of application usage in a week

As regards the remarks collected by the open ended questions one has to admit that they were quite significant. More specifically, the responses indicated that coaches who used the simulation had practical difficulties in using efficiently the application during the practice session. They had to interrupt their practice one or two times during a session, so as to go to the laptop running the simulation (which was placed on a court-side bench). On the contrary, the trainers with the iPhone devices incorporated the application quite well in their daily practice sessions. This was expected, as the smart phones are the platform of choice for deploying productivity tools addressed to professional trainers, who typically spend the most part of their work day without access to a typical computer. Thus, they can benefit from mobile applications that bring a lot of power in the convenient form factor of a mobile phone. Haptic interfaces employed in modern smart phones are an added bonus, having a lean learning curve and very intuitive operation.

Most users of the first group had difficulty in using the sinning dials of the time and date setting fields appearing on the new/edit drill, practice and program screens, some got confused with the monthly calendar view and tried to slide the month to the right in order to see the next month, while others thought that the (physical) button at the bottom of the device operates as a back (and not as an exit) and felt disoriented when they faced to the initial iPhone screen. In addition, novice iPhone users did not feel confident with the delete option of an item in a list (for instance when they attempted to delete an existing drill from a practice session). All those problems were not an issue for the experienced iPhone users.

Nevertheless, all coaches (even those that use the iPhone regularly) had some reservations as to whether they would be able to use the iPhone in the case of intense sunshine (will the screen contrast allow reading the text on the screen?), or what will happen in a case of a drizzle, or what if it is too cold and one has to wear gloves (the touch screen does not respond in this case). Other than those limitations that concern the weather conditions coaches have to work in, there were certain remarks concerning the difficulty of typing text using the iPhone virtual keyboard.

Concerning future versions of the application many coaches stated that it would be nice if the application could also allow them to mark athletes as present in a practice session. Moreover, they suggested that it would be very efficient if they could be able to receive the daily pracice schedule on the iPhone from the 'headquorters' (master coach) before getting to the court because usually they are not the ones that decide on the drills of each practice session. So if the application could be connected to a central (desktop/web) application and be somehow synchronised this would minimise the typing effort on their part and allow the update of the application data from a more comfortable position (indoors, sitted, on a large display, via a keyboard and a mouse).

5 Conclusions

This paper investigated the integration of hand-held haptic devices in the daily work activities of tennis coaches by evaluating the usability of the implemented Tennis Coaching Assistant running on iPhone. The evaluation was conducted in two phases. In the initial phase, 30 users were divided into two groups (depending on whether they are experienced iPhone users or not) and both groups were given a set of 7 usage scenarios to execute and a specially designed questionnaire to fill-in. A total of 84% of all users successfully completed 6 out of 7 scenarios and the feedback on SUS questionnaire questions was positive. In the second phase, 7 of the users in the initial group participated in field observation sessions and used the application during actual tennis practice and were interviewed to report on drawbacks/limitations they came across, as well as suggestions for improvements. Users incorporated relatively easily the application in their practice sessions and stated that it was helpful to have access to their practice programs and drills on court. A useful extension of the application would be to be able to enter new data using a desktop/web application on a PC and automatically synchronize the iPhone data, so as to avoid the most crucial efficiency problem they encountered: data entry using the virtual iPhone keyboard.

References

1. iPhone Dev Center, iPhone Human Interface Guidelines, Apple (2010)
2. iPhone Dev Center: iPhone Human Interface Guidelines for Web Applications, Apple (2010)
3. iOS Reference Library: iOS Human Interface Guidelines, Apple (2011)
4. Parrish, C., Ellis, B., Peters, K., Siebert, J., Wilson, E.: iPhone User Interface. Apress (2010)
5. Haywood, A., Boguslawski, G.: “I Love My iPhone.. *But* There Are Certain Things That ‘Niggle’ Me”. In: Jacko, J.A. (ed.) HCI International 2009. LNCS, vol. 5610, pp. 421–430. Springer, Heidelberg (2009)
6. Nihalani, P., Mayrath, M.: Statistics I: Findings from Using an iPhone App in a Higher Education Course, GetYa Learn On (2010)
7. Norman, D., Nielsen, J.: Gestural Interfaces: A Step Backwards in Usability. In: Interactions, vol. 17(5) (September / October 2010)
8. Piper, I.: Learn Xcode Tools for Mac OS X and iPhone Development. Apress (2009)
9. Barnard, D., Bondo, J., Burcaw, D., Kaneda, D., Kemper, C., Novikoff, T., Dalrymple, M., Knaster, S.: Learn Objective-C on the Mac. Apress (2009)
10. Mark, D., LaMarche, J.: Beginning iPhone Development: Exploring the iPhone SDK. Apress (2009)
11. iPhoneDev Center (n.d.), iOS Cocoa Touch, Apple, `http://developer.apple.com/technologies/ios/cocoa-touch.html`
12. Rosenberg, D., Stephens, M.: Use Case Driven Object Modeling with UML: Theory and Practice. Apress (2007)
13. Brooke, J.: SUS: a "quick and dirty" usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, A.L. (eds.) Usability Evaluation in Industry. Taylor and Francis, London (1996)

Mobile Web Usability: Developing Guidelines for Mobile Web via Smart Phones

Sungmoo Hong and Sung Chul Kim

KT Device R&D Center,
17 Woomyeon, Seocho, Seoul, South Korea
{sean,dkssc05}@kt.com

Abstract. More people have their smart mobile devices, more they access web sites via those devices. However mobile web users are not satisfied with experience of mobile web. What are the problems that users confront as they use mobile web? What are the design patterns of existing mobile web sites and what could be the recommended pattern among them? In order to answer these main questions, three different studies have executed: 1) task analysis, 2) in-depth interview, and 3) content analysis. According to findings from the studies, design guidelines for mobile web were finally established.

Keywords: Mobile, web, design, design guideline, smart phone.

1 Introduction

Data explosion is defined that the amount of data transferred among people will go much higher year after year so that it will explode. Cisco [1] predicted in their report that global mobile data traffic will "double every year through 2014, increasing 39 times between 2009 and 2014" reaching 3.6 exabytes per month by 2014. No one can doubt that smart mobile devices trigger this explosion.

Not only on data explosion but also on real life of mobile users, rapid evolution into smart devices has affected seriously and radically. It changes people's attitude and behavior toward web as well. More people have their smart devices, more they access web sites via those devices. Does it make any difference between access via PC and via mobile to web users? What are the problems that users confront as they use mobile web and how can we solve those problems?

1.1 Mobile Web Design

As most of mobile users were using internet via featured phones which had been built on relatively primitive operating systems by individual manufacture, they experienced problems which hindered their immersive experience of using mobile web: small screen size, abstract input devices (e.g., 12-keypad), throughput of mobile network, and limited interoperability between web and mobile web (in Korea, they had to use different mobile web, Wireless Markup Language (WML) and Wireless Application Protocol (WAP), which varied according to service provider and did not let them use generic full-sized web.)

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 564–572, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Smart mobile devices recently introduced to Korean mobile users (e.g., Apple's iPhone, HTC's Desire, and Samsung's Galaxy S etc.) seem to have been removing these barriers: they have big screen which is moderate enough to maintain mobility and haptic interfaces with touch screens, support broadband network (3G and Wi-Fi as well), and do fully support not only mobile web but also full-sized web. (Although iPhone does not support Adobe flash and Microsoft Active-X so that some features of web sites do not work, user experience of web surfing via iPhone has great advantages compared to that via featured phone.)

However, enjoying full-sized web sites via mobile phones is still a problem to some of mobile users. Nielsen [8] suggested in his column that “website use on mobile devices got very low scores, especially when users accessed "full" sites that weren't designed for mobile.” He explored on four main usability hurdles: small screens, awkward input, download delays, and mis-designed sites. Although in order to solve these problems many web sites providers are establishing mobile web adjusted to mobile screen size, network speed, and touch interfaces, it is hard to find specific guidelines or disciplines for mobile web, which is accessible via smart phones.

Even a decade before Nielsen, Jones et al. [4] showed an interesting result regarding the relationship between screen size and task completion rates of use of web sites. “The large screen group answered twice as many questions correctly than the small screen group.” Furthermore effectiveness of navigation on small screen was much lower than that on large screen.

Not only for users but also for designers mobile web is a challenge. Jones et al. [9] explained that “web designers are challenged even further by moving from the large screens and familiar input devices of the desktop computer, to the small, pocket-sized screens and limited interaction techniques of mobile devices.”

It is true that many researches and development have been done for improvement of usability and design of *full-sized* web. It is also true that improvement of *mobile* web could be rarely found [7]. This study poses research questions on mobile web and its usability regarding design.

2 Research Questions

In the context of mobile user experience mentioned above, KT, the biggest telecommunication company in Korea, faced serious demands from users who had complained their unhappy experience on surfing full-sized web sites via smart phones and from many content providers who had offered mobile sites fit to mobile browser provided by KT according to existing guidelines for developers and designers.

The purpose of this study is to revise pre-existing guidelines (titled “Design and Development Guidelines for Mobile Web Sites via KT Mobile Browser”) and expand its usage with more general purpose according to recent demands of great user experience of web surfing via mobile handsets.

Research questions are as below.

- What are the problems that users confront as they use mobile web?
- What are the design patterns of existing mobile web sites and what could be the recommended pattern among them?
- What are the principles for designing mobile web?

In order to explore these questions, three different studies have been performed: 1) task analysis, 2) in-depth interview, and 3) content analysis.

3 Study 1

3.1 Methods

First study, a task analysis was designed to find out the problems which users experienced during their use of mobile web. Participants were asked to use mobile web and observed when and how often they made errors. Within-subject design was applied to the study and tasks were counterbalanced.

Participants. Eight participants were recruited for the study followed by three different categories (gender, occupation, and experience of mobile internet) because those variables were supposed to be controlled. The participants are as follow: 4 females and 4 males, 4 college students and 4 employees, and 4 lay users and 4 experienced users.

Stimulu.The participants were asked to use two different mobile devices: LG's KU9000 and Apple's iPhone3GS (Fig. 1). Each device has its own browser, former KT Unified Network (KUN) browser and latter Safari. They were asked to use KT's mobile web portal SHOW (http://m.show.co.kr) on both devices and sites.

KU9000 iPhone3GS m.show.com on Safari browser m.show.com on KUN browser

Fig. 1. Devices (LG KU9000 & Apple iPhone3GS) and web sites (m.SHOW.com)

Usability. In order to understand where and when participants experience difficulties or errors while performing tasks on a web site - because "it is critical to improving the design of a site" [9] – they were asked to perform four tasks as follow. Main missions of these tasks were information seeking, communication, and transaction which were major mobile web activity themes identified by Cui & Roto [2].

1. Please browse news services and read several articles. You may use search function.
2. Please browse music services and hear several sample clips. You may download or send ringtones.
3. Please check out today's weather.
4. Please check out newly released films.

During their task solving, thinking-aloud and video record (Fig. 2.) were performed for better observation on their behavior (see Kjeldskov et al.'s study [6] for review of task solving and thinking-aloud method).

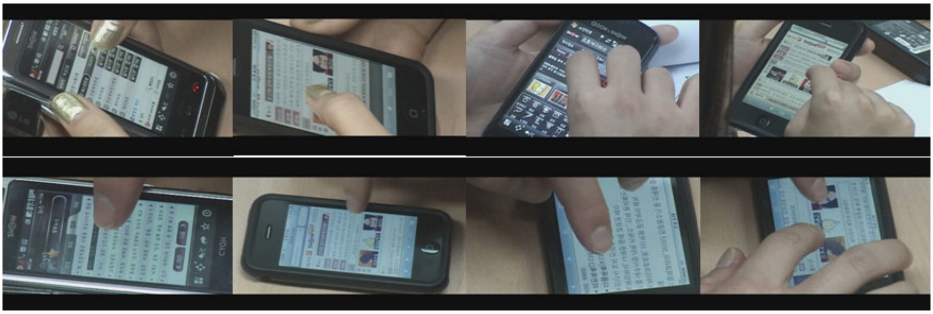


Fig. 2. Screen shots captured from video clips of participants' task-solving session

Procedure. Entering the laboratory, participants were given brief introduction about the research and then given ten minutes for making familiar with stimulus devices. Facilitator told them to perform four tasks described on prior Usability section one by one and thinking-aloud as well. He was not to guide them how to perform tasks successfully. The whole sessions were video-recorded digitally for analysis of patterns of errors and difficulties which they experienced. After task-solving session, an in-depth interview was followed (this is discussed on study 2 section.), thanked with reward of fifty dollar cash, and dismissed.

Results. By analyzing video clips and participant's comment of thinking-aloud, two main problems of usability of mobile web were revealed, 1) interaction, and 2) navigation error. Mostly they experienced difficulties as they interacted with touch interface because the component of graphic user interface (e.g., buttons, icons, and other text-hyperlinks) was too small to touch. Especially lay users performed this touch interaction error (M=4.5) more than experienced users did (M=2.5). They also confronted errors as they were navigating (or browsing) both within a page (scrolling up & down) and through various pages (before & after pages and up & down depth navigation). Lay users experienced difficulties by performing errors (M=3) more than experienced users did (M=1.5).

According to analysis of participants' behavior, most of them were using only one finger to interact with mobile devices and web sites. Especially female participants experience difficulties as they interact with one finger due to their relatively long fingernails. Based on "almighty one finger [5]," the participants reported during their thinking-aloud session that scrolling did not matter as they were using a browser modified with touch interface though it was one of the biggest problems of featured phone and its browser.

In sum, the size of objects which were displayed on small screen and supposed to be touched with one finger affected negatively on the usability of mobile web.

4 Study 2

4.1 Methods

Second study followed by the first study was performed in order to get some valuable qualitative insights to enhance design guidelines for mobile web. A semi-structured interview was applied to each of eight participants who had been participated to study 1.

Procedure. After they had performed the first study, task-solving session, they were given a pencil & paper based questionnaire sheet. They were asked to fill in basic demographic information and four evaluation questions, which asked to evaluate each of four tasks (e.g., “what made you difficult to check out today’s weather?”). Then interviewer interviewed participants based on their answers.

Results. Most of interviewees complained about complexity of web sites. They did not like web sites crowded with too many texts and too many objects to be touched. They blamed poor usability of text input on small screen with QWERTY-based-on-screen-touch keyboard. They wanted simple and big web pages and light depth of structure though they consequently elicited long down scrolling (Table 1.). In sum, participants did not like rich experience browsing web but simple experience on mobile web.

Table 1. Evaluation on experience of mobile web

Evaluation	Lay users	Experienced users
Overall	• Too complicated. • Too many texts. • Hard to figure out what the content is.	• Too small icons and buttons. • I want simple!
Web surfing	• Want simple and big. • Want one glance for one page. • Want less text. • Too many depths.	• Headlines are enough for main page. • Simple and big beats scrolling.
Comparison to keypad	• Hard to zoom in and out. • Text input is annoying.	• Annoying text input.

5 Study 3

5.1 Method

In order to analyze design patterns of existing mobile web sites, 33 mobile web sites were categorized according to their design layout as a content analysis. In order to elicit optimum size of GUI component grid system prototyping was performed.

Design pattern. Screen shots captured from representative pages of web sites were printed out and categorized according to their page layout and design by two design experts. (Fig. 3.)

Fig. 3. Analyzing design pattern of existing mobile web sites

Grid system prototyping. Grid patterns of 33 mobile websites were printed on vinyl pad (Fig. 4. (a)) in order to investigate cumulative trace of all of grid layouts. The purpose of this examination is to analyze efficiency of space according to page design and grid system and to find out proper size of GUI component. Special application which could show and save cumulative traces of touch interaction was used during the analysis (Fig.4. (c)). Two design experts tested followed by this process.

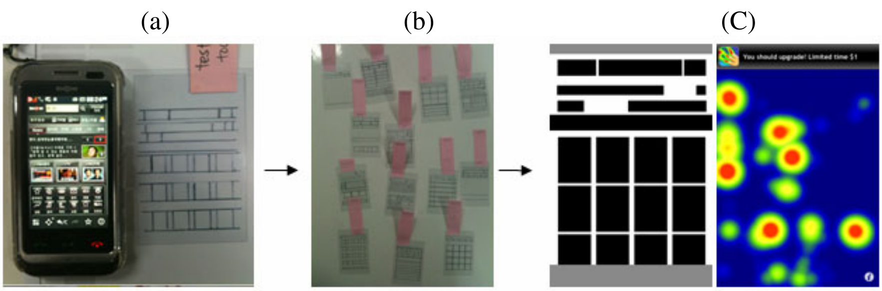


Fig. 4. Process of grid system analysis

5.2 Results

Analysis of design pattern elicited five types of standard design layout, 1) mixed type, 2) icon type, 3) image type, 4) text type, and 5) search type. (Fig. 5.) Optimum size of component which did not go over the size of each trace of touch interaction was elicited as 4.5mm (width) X 6.5mm (height) (Fig. 6.).

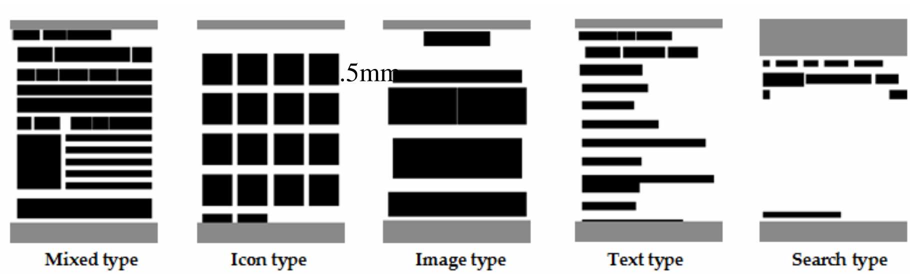


Fig. 5. Five types of design pattern of mobile web

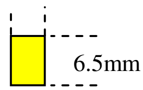


Fig. 6. Optimum size of touchable GUI component (actual size on MS Word)

6 Discussion and Application

Small screen of mobile device and unoptimized web sites distract users as they surf on web sites via mobile devices. In the study mobile web users performed errors as they interact with web sites due to small GUI components, complex layout and design of web pages, and complicated structure of web sites and annoyed poor text input interface. Findings of current study suggest that optimized mobile web site would be a must for service provider. However, “while there has been much successful work in developing rules to guide the design and implementation of interfaces for desktop machines and their applications, the design of mobile device interfaces is still relatively unexplored and unproven.”[3]

Current study was executed in order to revise existing guidelines (titled “Design and Development Guidelines for Mobile Web Sites via KT Mobile Browser”) and expand its usage with more general purpose according to recent demands of great user experience of web surfing via mobile handsets. As an application, newly revised guidelines were developed. The design guidelines (Korean version) are now available for downloading from KT developers sites. It is the first design guidelines for mobile web sites in Korea. Its main contents are as below.

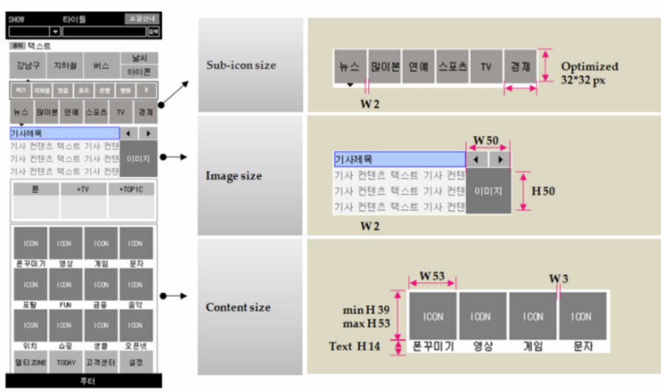


Fig. 7. Layout guideline (p.21)

1. Recommended layout guideline (Fig. 7)
   Basic guidelines for layout are identified in the revised guidelines. In order to meet user’s needs of immediate information-seeking, search box should be on top of each page. Layout design is in the full range of freedom for content designer and developer other than pixel size of each component of content (e.g., main and sub icon, image, and text) in order to guarantee errorless touch interaction.

2. Five types of sample pages according to recommended layout (Fig. 8)
   In order to help designers to design web pages more rapidly, the revised guidelines conveyed five types of sample pages of representative web sites: icon, image, text, search, and mixed type. Sample codes are also appeared in appendix section.
3. Library of GUI components (Fig. 9)
   Recommended GUI components are delivered with guideline document as GUI library. They are to be used frequently or mandatorily, e.g., button of search, price, more, before, next, numeric, log-in, download, help, and input box.

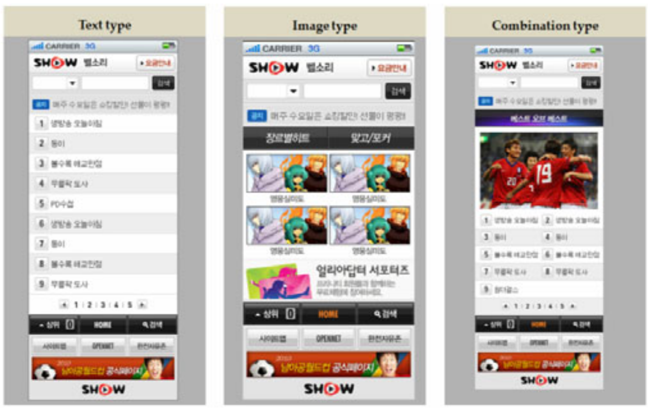


Fig. 8. Sample pages in regard of type of layout (p.23)

Fig. 9. Library of GUI components

7 Limitations

Although main purpose of the study was not a quantitative investigation, more sample participants would elicit more solid and specific direction for establishing design guidelines. Iterative design process (e.g., prototyping - user test - redesign process) would also benefit better design output of guidelines.

Above all, common problem of this kind of process of guideline and standardization is to be a mandatory. In the previous stage of mobile web, content provider should pass a test by operators because they use operator's browser and pricing protocol. However, in this smart mobile web stage, they do not need to follow operator's guideline mandatorily.

Mobile web services are supposed to be an alternative or ultimate future of recent mobile application services. In order to offer great experience to mobile users during their use of mobile web, every designer and developer should keep in mind that mobile and PC web are different.

References

1. Cisco Systems, Inc. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update (2010), http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html
2. Cui, Y., Roto, V.: How People Use the Web on Mobile Devices. In: WWW 2008 / Alternate Track: Industrial Practice and Experience, pp. 905–914 (2008)
3. Gong, J., Tarasewich, P.: Guidelines for handheld mobile device interface design. In: Proceedings of DSI 2004 Annual Meeting (2004)
4. Jones, M., Marsden, G., Mohd-Nasir, N., Boone, K., Buchanan, G.: Improving Web interaction on small displays. Computer Networks 31, 1129–1137 (1999)
5. Kim, H., Kim, J., Lee, Y., Chae, M., Choi, Y.: An Empirical Study of the Use Contexts and Usability Problems in Mobile Internet. In: Proceedings of the 35th Hawaii International Conference on System Sciences (2002)
6. Kjeldskov, J., Skov, M.B., Stage, J.: Instant Data Analysis: Conducting Usability Evaluations in a Day. In: Proceedings of the third Nordic conference on Human-computer interaction (2004)
7. Koskinen, E., Kaaresoja, T., Laitinen, P.: Feel-Good Touch: Finding the Most Pleasant Tactile Feedback for a Mobile Touch Screen Button. In: Proceedings of the 10th international conference on Multimodal interfaces (2008)
8. Nielsen, J.: Mobile Usability. Jakob Nielsen's Alert Box (2009), http://www.useit.com/alertbox/mobile-usability.html
9. Waterson, S., Landay, J. A., Matthews, T.: In the Lab and Out in the Wild: Remote Web Usability Testing for Mobile Devices. In: CHI 2002 extended abstracts on Human factors in computing systems (2002)

Usability Analysis in Gesture Operation of Interactive E-Books on Mobile Devices

Ching-Hua Huang and Chao-Ming Wang

Graduate School of Design Doctoral Program, National Yunlin University of Science and Technology, 64002Yunlin, Taiwan
{g9930818,wangcm}@yuntech.edu.tw

Abstract. This paper analyses usability of interface including gesture operation and interaction in HCI (human-computer interaction). We explore the common gestures which are used to read e-books, and summarize the mode of operation of these gestures. This analysis bases on the five indicators that are proposed by Nielsen in 1993 to assess user interface of the interaction e-book. It discusses the interface of e-books that experts read and places emphasis on key points for visibility, button design and operation mode of multimedia button. Heuristic evaluation was adopted to study the gestures operation in reading e-book modes based on mobile devices. The study concludes that operation mode, perception & feedback design, and user's behavior are important interactions mode for interface design on mobile devices. The paper offers interaction design principles for development and improvement of the interactive e-book.

Keywords: Interactive e-Book, Gesture Operation, Usability Analysis, Mobile Device.

1 Introduction

The advance in electronic technology leads the trend and the mode of digital learning. Due to the development of internet the publishing industry experiences the power of digital knowledge transmitted over the network. They saw the limitations of traditional publishing, and realized e-book as the trend of modern learning. Approximately 5.9 million U.S. adults own an eReader according to the latest data from GfK MRI--up from 2.1 million owners in the March-October period of 2009 when GfK MRI first began asking consumers about their usage of devices like the Amazon Kindle and SONY Reader. In the period of survey of March-October 2010, 49% of eReader owners were male and 51% were women. According to the survey, currently 55% of eReader owners have annual household incomes in excess of USD$100,000[1].

Since mobile devices of which production owes to new media and new technology are endowed with the advantage of flexibility, and since user behavior has changed, touch screen is widely used on mobile devices, such as iPhone, iPad and HTC. These mobile devices use multi-touch control methods for intuitive operations, such as drag and zoom, which promote the trend of e-books, and built-in auto-rotation, intelligent sensing and other new technologies.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 573–582, 2011.
© Springer-Verlag Berlin Heidelberg 2011

This study discussed the mode of gesture operation used to read e-books on mobile devices, sorted out the gestures common operations on the touch screen to read e-books by the literature, analyzed the relation between the user's perception and touch screen user interface design to find out the relationship between efficiency and error rate.

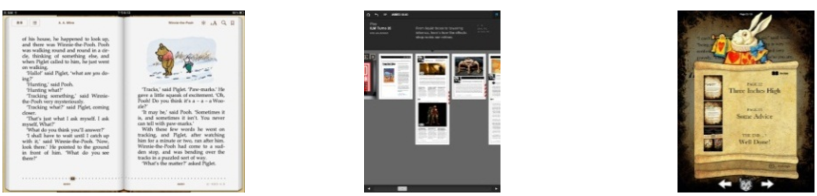
The purposes of this study are as follows:

1. To explore the common gesture which has been used to read e-books, and to summarize the mode of operation commonly used gestures.
2. To analyzed the relationship between sensors on mobile devices and interactive books.
3. To summarize the principles of interactive design interaction for the e-books on mobile devices.
4. To assess the interaction e-book user interface according to five indicators, proposed by Nielsen in 1993[2].
5. To propose features and differentiate services that the e-books on mobile devices may have.

2 Literature Review

2.1 A Study of Present E-Book

The e-book is a new information technology product that facilitates reading and acquisition of information. It is a written work readable on the screen of a PC, a PDA (personal digital assistant), or a reader specifically designed for the purpose. It provides the same meaning as a conventional paper book (C-book) which can stores and communicate knowledge through reading. On the positive aspect, an E-book is superior to a C-book from diverse perspectives such as storage, transfer, delivery, and accessibility[3]. Recently, many types of e-book readers (see Table.1) have been developed for iPad device as (Fig.1):



(a) iBooks (b) WIRED magazine (c) Atomic Antelope

Fig. 1. Types of e-books

iBooks. The iBooks app is free download form app store. Readers can select and download the e-book to their libraries, or import PDF files to the iBooks. Readers can use gesture operation to move from page to page on the e-book. Additionally, readers can also select and include interactive page ctrl option, which let readers drag a page corner to turn the page[4].

WIRED. WIRED has approved the app as an interactive digital edition of the magazine. The WIRED app is available from the iPad app store. Every page in the issue is individually designed for optimal viewing on the iPad screen in both portrait or landscape orientation. It is an electronic document that supports a variety of interactive function on the touch screen. Readers can open a new page or change dynamic mode by selecting a button on the e-book. Readers can also use the action buttons to get multimedia content[5].

Atomic Antelope. Alice for the iPad by Atomic Antelope is an interesting book which includes game, sensor, illustration and story book. By providing some of breathtaking animated scenes, it revolutionizes the Lewis Carroll classic. Pictures that move and animate as you tilt the device. The interactive e-books were entertaining kids on car journey[6].

Table 1. Elements of visual analysis on e-books

Sample	Description	Element
iBooks	The iBooks content which is the most important thing is word and the second one is picture. The iBooks App provides basic function (e.g., scale, label…) to read.	Word/ Picture
WIRED	The WIRED magazine combined with a variety of multimedia elements (e.g., video, audio, VR).	Word/ Picture Video/Audio 360 VR Button
Atomic Antelope	The Alice for the iPad which was sensor and multi-touch features to create interesting interactive situation.	Word /Picture illustration animation

2.2 Interaction Mode Gesture Operation on Mobile Devices

A sensor is typically an electrical or electronic component whose job is to detect changes in the environment. These changes can be any number of things, depending on the type of sensor. The most common types currently used for interactive gestures include pressure, light, proximity, acoustic, tilt, motion and orientation [7].

Currently, most gestural interfaces can be categorized as either touchscreen or free-form. Touchscreen gestural interfaces—or, as some call them, touch user interfaces (TUIs)—require the user to be touching the device directly. For example, these would typically include(Fig.2)[8]:

1. Tap: Briefly touch surface with fingertip
2. Double tap: Rapidly touch surface twice with fingertip
3. Flick: Quickly brush surface with fingertip
4. Drag: Move fingertip over surface without losing contact
5. Pinch: Touch surface with two fingers and bring them closer together
6. Spread: Touch surface with two fingers and move them apart

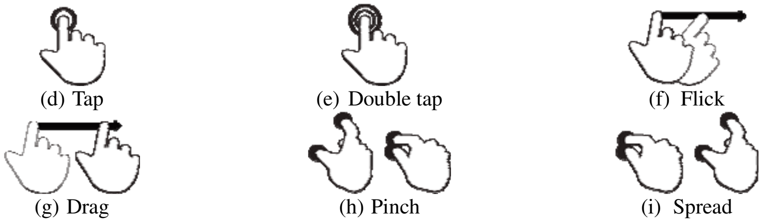


Fig. 2. Gesture operation on touch screen

2.3 Usability Evaluation for Mobile Devices

Nielsen indicates that a variety of methods e.g., Heuristic evaluation, Performance measures, Thinking aloud, Observation, Questionnaires, Interviews, Focus groups, Logging actual use, and User feedback, can be applied for usability evaluation. For the reason of space, it is necessarily simplified, but it still provides a good, quick overview of these methods. It is obvious that these methods are intended to supplement with each other, since they address different parts of the usability engineering lifecycle, and since their advantages and disadvantages can partly make up for each other. It is therefore highly recommended not to rely on a single usability method to the excluding others[2]. Also, Nielsen indicates that the choice of method may be partly dependent on the number of users that are available for usability activities. If very few users are available, emphasis should be placed on heuristic evaluation, thinking aloud, and observation.

There are many possible ways to combine various usability methods; each new project may need a slightly different combination; it depends on its exact characteristics. A combination that is often useful is that of heuristic evaluation and thinking aloud or other forms of user testing. Typically, one would first perform a heuristic evaluation to clean up the interface and remove as many "obvious" usability problems as possible. After a redesign of the interface, it would be subjected to user testing both to check the outcome of the iterative design step and to find remaining usability problems that were not picked up by the heuristic evaluation.

This study analyzes the current situation according to the above e-book, and summarizes the characteristics of interface design should have in Table 2.

Table 2. The characteristics of interface design

Items	Functions	Characteristics
Multimedia buttons	Video control Audio control Pinch and zoom Scroll	Striking tool prompt Intuitive button design
Navigation mode	Story view Page switch	Quick page switching mode Clear navigation map
Content menu	Slide show	Suitable menu prompt Appropriate guidance reading

3 Methods

Heuristic evaluation is a discount usability engineering method for quick, cheap, and easy evaluation of a user interface design. Heuristic evaluation is the most popular method of usability inspection. Heuristic evaluation is done as a systematic inspection of a user interface design for usability. The goal of heuristic evaluation is to find the usability problems in the design so that they can be attended to as part of an iterative design process. Heuristic evaluation involves a small set of evaluators that examine the interface and judge its compliance with recognized usability principles (the "heuristics ")[9]. Therefore, this study is intended to search the interface usability experts, which have conducted the heuristic evaluation.

3.1 Participants and Tools

Single evaluators found only 35% of the usability problems in the interfaces. However, since different evaluators tend to find different problems, it is possible to achieve substantially better performance by aggregating the evaluations from several evaluators. There is a nice payoff from using more than one evaluator, and it would seem reasonable to recommend the use of about five evaluators, and certainly at least three[2]. Therefore, this study has selected five interface usability experts, such as interface usability, interface design and programming to evaluate interactive interface design for WIRED magazine on the iPad. Finally, researcher was summarized reader gesture in interactive e-book of interface usability principles on mobile devices.

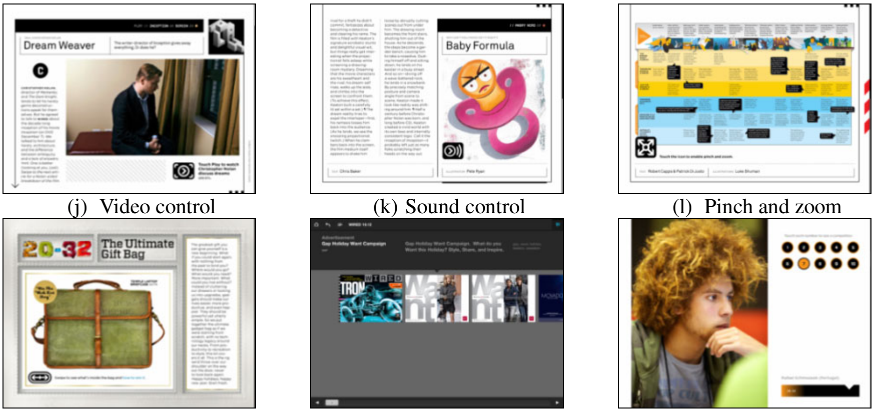
3.2 Procedure

The researchers conclude based on the literature of interactive e-books should have the interface design features: multimedia buttons, navigation mode, content menu. First, listed the main page of interactive design on WIRED magazine, and then according to Nielsen in 1993 proposed the principle of the five criteria (learnability, efficiency, memorability, errors, satisfaction), the last question of usability. The steps were stated as follows:

1. Inviting five experts evaluation of WIRED Magazine.
2. Explaining the process of heuristic evaluation.
3. Sending to the expert evaluation questions of WIRED Magazine.
4. Conducting for each expert evaluation of 1-2 hours.
5. Communicating comments with each other, and concluding them in the discovery of interface usability.
6. Integrating evaluation results and recommendations.

3.3 Usability Evaluation

Fig.3 shows the common page when user read the WIRED magazine on the iPad. Expert evaluation might take about 1-2 hours. The researcher completely recorded the operating gestures and error rate of interface usability problems in the interactive e-book. Finally, we arranged the production features for the e-books on the iPad.



(j) Video control (k) Sound control (l) Pinch and zoom

(m) Scroll (n) Navigation (o) Content switch

Fig. 3. Features of WIRED magazine's interface design

Focus on recording and analysis:

1. Analyzing the advantages and disadvantages of WIRED magazine interface?
2. Collecting common gesture used to read e-books, in addition to the above mentioned documents outside of basic gestures, what?
3. Using five criteria which were proposed by Nielsen in 1993 as an index to evaluate WIRED magazine's user interface.
4. Analyzing the relevance between sensors of existing mobile devices, and interaction e-books.
5. Archiving interactive design principles of interactive e-books of mobile device.

4 Research Results and Discussion

4.1 Results of Heuristic Evaluation

We organized practical user evaluations within the setup of a heuristic evaluation by experts. Five experts (four males and one female) were invited. All of these experts have experience in multi-touch panel and three of them are regular users of iPad. This study investigates the background of these experts (Table 3) and usability of interface design and operation mode on WIRED magazine.

Table 3. Background information of the subjects

participant number	1	2	3	4	5	Average
Sex	M	M	F	M	M	4M, 1F
Age	32	27	30	33	32	30.8
Years of work experience	2	1	4	7	5	3.8
Experience with e-books	Y	Y	Y	Y	Y	5Y
Experience with multi-touch panel	Y	Y	Y	Y	Y	5Y
Experience with iPad	Y	Y	Y	N	N	3Y, 2N

These heuristics were given to five experts to evaluate WIRED magazine on iPad independently. The researchers recorded all the operating gestures and problems of interface usability that came forth in the evaluation.

Table 4 shows the evaluation results of expert reading e-books on three aspects: key points for visibility, button design and operation mode of multimedia button.

Table 4. Key points of heuristic evaluation results

Function	Icon	Gesture	Visibility	Button design	Operation mode
Video control		Tab	Passably visible, but easily be taken as background	Simple. recommended to animate button	Easy to operate
Audio control		Tab	Easily overlooked	Easy to be confused with the video control	Simple operation. Advisable to show time line.
Pinch and zoom		Tab, Pinch, Spread	Prompt was not clear	Image design disagreed with the user cognitive	Zoom could not be as free as photo operation
Scroll		Flick, Drag	Easily overlooked	Image design easy to cause misunderstanding	Simple operation, but prompt was not clear

Summaries of the key aspects of each evaluation paradigm for the following issues:

1. Multimedia buttons: WIRED magazine contained a lot of multimedia elements. User could start button when he wanted to read more contents. Multimedia buttons were easy and quick operation, the most common gestures were tap and double tap.
   - Visibility: Content of the WIRED magazine's video could passably be identified, but some of the control buttons are not obvious.
   - Button design: Video control button is easy to be confused with the audio control buttons, pinch and zoom button disagreed with the user's cognitive. Button design recommends animated buttons in order to attract user’s attention.
   - Operation mode: It takes time for beginner to become familiar with the touch screen. It is propose to show time line and provide functionality for adjusting volume while sound is played. Using pinch & spread could scale image, but not as free as in and out photos zoom. Images drag button could not express operation mode clearly, it is proposed to use tap and prompt with the direction of arrow.
2. Navigation mode: Navigation interface of WIRED magazine could be quickly and easily used. But for beginner it needed time to find the navigation, and the icon design was not easy to understand. The navigation scroll underneath the screen facilitated quick access to the designated page, but there was the problem of turning too many pages. The commonest gesture for turning pages was flick.
3. Content menu: The function was simple and easy to operate, the common gesture for switching content was tap, but the color on the design of the menu was inconsistent, because some were number buttons, and others were image buttons, so it requires care to look out for. In addition, users jumped easily to other pages by touching out of mistake, and found it difficult to recover.

Some of the features that could be improved include: learnability, efficiency, memorability, errors, satisfaction. It should be taken note that issues were ranked in ascending priority: low, medium, and high. Table 5 shows improvements and recommendation suggested by experts.

Table 5. Recommended improvement of the result of heuristic evaluation

NO	Issue Priority	Issue	Recommendation
1	Medium	For the first time young men could quickly get started, but older men without the guidance found it difficult.	Advisable to add help pages or guidelines that enable users a quick start.
2	High	The image button was easy to be confused with the content, and difficult to be identified.	Provide unified image button and create animate button for highlighting.
3	Medium	Magazine with thematic modes, among which some flipped left or right and the other flipped up or down. This causes confusion.	Provide choice by addition of two-page mode.
4	Low	Contents could be adjusted vertically or horizontally to read with sensor, but some media would not display in certain type.	Several tests are needed to confirm that the e-book can play and read in both vertical and horizontal type.
5	High	Attention is not easily drawn to navigation buttons.	Navigation buttons should be made more prominent.
6	High	If user had never used it, he did not know how to use the image zoom	Addition of help pages to guide the user for a quick start.
7	Low	It was difficult to restore images when they had been enlarged and causing the turn page unsmooth.	Addition of undo key for quick undo.
8	Medium	Not all images and content could use the zoom function.	All images should contain as much as possible the facility of zoom function.
9	High	Articles with different lengths and mixed with ads are not provided with clear prompt in reading. It is easy to make mistake for searching reading page.	Ads and article should be clearly marked so that user could make choices.
10	Medium	It is easy to lose the position while browsing content.	Addition of chapters page.
11	Low	Fonts too small, not easily readable for everyone.	Provide possibility of changing fonts. Font colors should be made distinct from the background color.

4.2 Satisfaction Survey Results

This study investigated the expert's usability of interface design and operation mode on WIRED magazine. Table 6 show results of usability satisfaction.

Table 6. Results of usability satisfaction

Question	Very unsatisfied	Unsatisfied	Neutral	Satisfied	Very Satisfied
Learnability			1	4	
Efficiency		1		4	
Memorability		1		2	2
Errors		2	3		
Satisfaction		1		4	

The results showed that experts felt satisfied with the learnability, efficiency, memorability, errors and satisfaction of the using of WIRED magazine. WIRED magazine classified content and advertising separately, but the classification was not clear leading difficulty in reading and in the usage of interface. Therefore, e-book should contain clear illustrations to guide readers. It should provide some interactive information to give readers a sense of participation and rich content for reading. The study concludes interaction design principle of interactive e-books for mobile devices as follows:

1. Easy usage
2. Guided reading
3. Clear illustrations
4. Interactive mode
5. Sense of participation

5 Conclusion

This study provides usability of reading e-books for mobile devices. These heuristics were given to five experts who have evaluated WIRED magazine on iPad independently. We compiled their comments and discussed with them about their findings in a meeting and obtained suggestions from them on strategies for addressing problems. The following are some important points of their findings:

1. Learnability: Help pages or clear guidelines should be added for a quick start for beginner.
2. Efficiency: Navigation buttons should be made more prominent for easy usage and more satisfaction of user.
3. Memorability: Gestures commonly used for the touch-screen should be found out for simple usage to read e-book for WIRED magazine.
4. Errors: More tests are needed to confirm readability in vertical and horizontal type.
5. Satisfaction: Ads and article should be marked clearly for easy choices of user.

The study concludes that operation mode, perception & feedback design, and user's behavior are important interaction mode for interface design on mobile devices. This study has shown the interaction design principles (easy usage, guided reading, clear illustrations, interactive mode and sense of participation) for development and improvement of the interactive e-book.

1. Operation mode: Addition of multi-function sensors (e.g., pressure, light, proximity, acoustic, tilt, motion, orientation) in order to design the service interface
2. Perception & feedback design: High visibility of button to guide user, and give real time feedback to enable usability satisfaction.
3. User's behavior: Tap, double tap, flick, and drag were common gesture for reading e-books, and intuitive gestures could improve user satisfaction. Pinch and spread gestures should be added to complete instructions so that users could quickly get started.

This heuristic evaluation which involves a small set of evaluators to examine the interface is in accord with Nielsen's suggestion. We have to acknowledge the limitations with a sample size of five in this study. In order to achieve a higher validity we hope to have a larger survey sample for the future research.

References

1. GfK MRI. eReader Ownership By Adults Nearly Triples in Less Than Two Years. GfK MRI Survey of the American Consumer (2010)
2. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
3. Press, L.: Personal computing: from P-books to E-books. Commun. ACM 43, 17–21 (2000)
4. Apple Inc., `http://www.apple.com/itunes/`
5. Tron's Triumphant Return to Cyberspace. In: WIRED, 18.12. Condé Nast Publications, United States (2010)
6. Antelope, A.: Alice for the iPad. In: Antelope, A. (ed.) Atomic Antelope, London (2010)
7. Saffer, D.: Designing Gestural Interfaces. O'Reilly Media, Inc., Sebastopol (2008)
8. Luke Wroblewski, `http://www.lukew.com/touch/`
9. useit.com. Jakob Nielsen's Website, `http://www.useit.com/`

Mobile QWERTY User Research

Yat-Sang Hung[1], Sarena Yang[1], Stephanie Vance[1], and Neung Eun Kang[2]

[1] Samsung Telecommunication America, 1301 E. Lookout Drive,
Richardson TX 75082, USA
[2] Samsung Electronics Ltd., 10F, Samsung Electronics, 1320-10 Seocho 2-dong, Seocho-gu,
137-857, Seoul, Korea
{sang.hung,sarena.yang,s.vance,ne.kang}@samsung.com

Abstract. In this paper, we describe the methodology and the activities of mobile QWERTY research and give examples of different pieces of research and some of the results and insights. To minimize confounding effects of different aspects of QWERTY design, we use virtual keyboard layout tools to construct/replicate different keyboard layout designs so that both human performance and user satisfaction can be captured by systematically varying design parameters in the same testing environment. In addition, we also conduct usability testing by varying specific hardware design parameters and keeping all others constant whenever possible. By using this 2-pronged approach, an empirical performance model of mobile QWERTY keyboard in relation to different design attributes (both perceptual and mechanical) can be built over time. The approach we have adopted in virtual keyboard testing also helps to improve virtual QWERTY keyboard design on touch devices as well as physical QWERTY keyboard layout on devices with physical QWERTY.

Keywords: mobile QWERTY, human performance, user satisfaction, performance model.

1 Introduction

Usability of mobile QWERTY keyboard is determined by many different design variables: the overall size of the keyboard, the number of lines, the size, height, trip force, travel of individual keys, the geometry of key shape etc. Samsung Design Group at Korea and Samsung Telecommunications America jointly conducted systematic QWERTY ergonomics and usability evaluations of existing as well as prototype mobile devices to unravel the relationship between different design variables and establish best practices of mobile QWERTY keyboard design.

1.1 Research Goals

The primary goal in this program is to guide the development of new QWERTY devices by best practices informed by research and to test and improve new QWERTY keyboard design prototypes in development. These include the design and development of both physical and virtual QWERTY keyboards. By benchmarking typing performance, subjective user satisfaction and assessment of usability, we will continually improve our design recommendations and best practices.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 583–592, 2011.
© Springer-Verlag Berlin Heidelberg 2011

1.2 Research Methodology

In order to provide research supporting a wide spectrum of situations, different research techniques have been used, ranging from low-fidelity paper prototyping using simulated typing to relatively large scale competitive text entry with different QWERTY devices in the market. In addition, competitive text entry research using virtual QWERTY keyboards has been conducted routinely to compare different QWERTY layouts from different device vendors or carriers and keyboards from third-party developers.

1.3 Measurements

In most cases, dependent measures included human performance (speed and accuracy of typing), user satisfaction, and evaluation of the ease of entering text, symbols, and numbers. Within-subjects counter-balanced experimental designs have been used when possible in most of the research studies conducted. In all the research done, the branding information of the devices or prototypes was hidden to avoid potential bias of the respondents.

The speed of typing was captured by having respondents type a template string as fast and accurately as possible within one minute without correcting any mistakes. The speed was measured by the number of characters typed per minute (cpm) including spaces, numbers and symbols.

The accuracy of typing was derived from the mean string distance (Levenshtein distance) between the typed string and the template string [1].A composite performance index (reflecting both speed and errors) was defined from the relative percentage of correct matches with the template string.

Subjective assessments of ease of text, number and symbols entry, clarity of marking on the QWERTY keys and overall typing satisfaction were recorded on 1-5 Likert scales. The devices under comparison were also ranked in terms of overall QWERTY design preference by the respondents.

In the study of the relationship between performance, user satisfaction and physical design characteristics, the following QWERTY keyboard design characteristics were measured: key size (height, width), key travel, trip force, keyboard width, keyboard height, center-to-center key distance (both width and height), number of key rows, and the space bar width. Subjective assessment of the physical design characteristics were captured by a 7-point scale with “too small” on the left, “just right” in the middle (at 4) and “too big/large” on the right.

2 Physical Mobile QWERTY Evaluations

From these subjective evaluations as well as objective measurements of human performance, the optimum range of physical design characteristics can be determined. The primary research activities in this area include competitive performance benchmarking of devices with physical QWERTY keyboard, collection of qualitative feedback and ratings with mockups and wax models.

The various physical design parameters (typical measurements include trip forces, key travel, key sizes, spacing between keys etc.) are measured and correlated with both performance and subjective evaluation of the different physical design characteristics. Example research in this area includes an investigation of the relative roles of different physical parameters using mobile physical QWERTY devices in the market with varying design characteristics, comparison of key trip force by using two identical keyboards with different trip forces, and comparison of key depth (the height of the key from the adjacent surface) using wax models.

2.1 Competitive Mobile Physical QWERTY Performance Benchmarking

Competitive mobile physical QWERTY performance benchmarking studies were conducted frequently on representative devices in the market as well as Samsung pre-launch devices. Typically, participants would perform typing tasks on a number of devices with physical QWERTY keyboard. The order of the devices would be counterbalanced and the respondents would type 1 and/or 2 predefined strings (one would consist of mostly words while another would be a mix of words, number and symbols). They were instructed to type the string as fast and as accurately as possible within one minute without correcting any mistakes. The results were analyzed against the template strings. The number of characters typed was counted and the mean string distance from the template strings was calculated.

The results from a recent QWERTY performance and user satisfaction of LG Ally, HTC G2 and Samsung Epic Android smartphones is shown below [2].

Respondents (N=54) completed a typing task with mixed input entry (letters/symbols/number) using LG Ally, HTC G2, and Samsung Epic. Both LG Ally and Samsung Epic keyboards were ranked significantly higher than the keyboard of HTC G2. The same significant results applied to overall satisfaction. respondents also typed significantly faster and more accurately with both LG Ally and Samsung Epic Keyboards, and the key markings of Samsung Epic were rated significantly better in clarity than the other devices.

HTC G2 was rated significantly worse than other devices in the ease of typing symbols, the ease of text entry and the ease of number entry.

Table 1. Competitive QWERTY performance and usability scores

QWERTY Benchmark	LG Ally (A)	HTC G2 (B)	Samsung Epic (C)
Speed (cpm)	49.29*	45.02	52.15*
Accuracy (	0.34*	0.30	0.36*
Ease Text (1-5)	3.52*	3.04	3.69*
Ease Symbols (1-5)	3.13*	2.37	3.28*
Ease Numbers (1-5)	3.93*	2.78	4.13*
Clarity (1-5)	3.91	3.22	4.06*
Satisfaction (1-5)	3.52*	2.81	3.41*

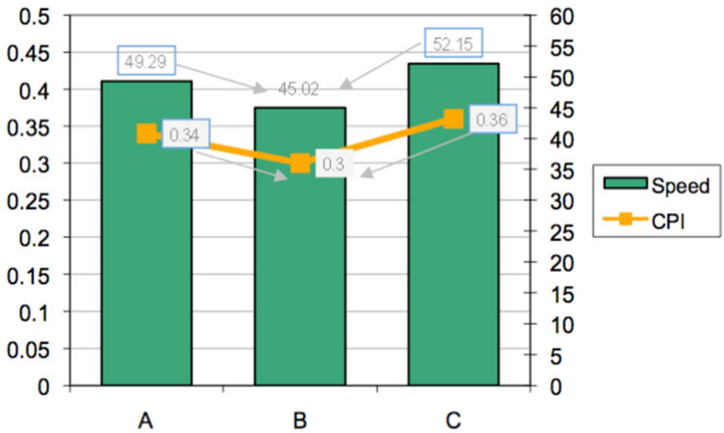


Fig. 1. Speed and Composite Performance Index of LG Ally (A), HTC G2 (B) and Samsung Epic (C)

2.2 Physical Design Attributes and QWERTY Performance/Satisfaction

The primary purpose of this research is to build a performance/user satisfaction model of QWERTY based on salient physical design characteristics. This line of work also spurred interests from the carriers on the usability of QWERTY keyboards and their own investigation of QWERTY comparisons.

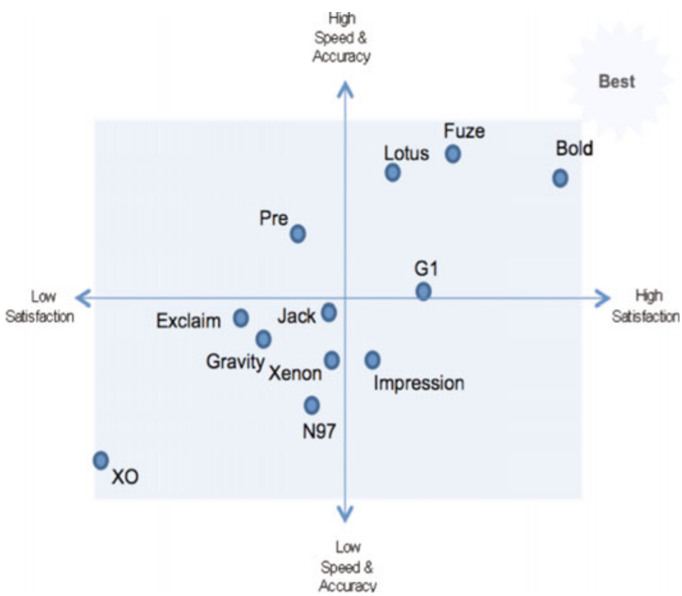


Fig. 2. Speed/accuracy and satisfaction of 12 different physical QWERTY devices

The following example study [3] was used to develop a set of design parameter recommendations and guidelines. A QWERTY performance and user satisfaction study (N=119) was conducted with 12 different mobile phones with physical QWERTY keyboards. The physical design characteristics of each keyboard were measured. The performance data (speed and accuracy) and user satisfaction data were used to identify critical physical design characteristics of mobile QWERTY keyboard.

The baseline analysis was a conservative Partial Linear Square Regression on 12 QWERTY devices (2 Components were used for model quality) in this study:

1. Dependent Variables: Speed, Accuracy, Ease of Typing and Satisfaction.
2. Independent Variables: The following baseline physical design variables have been identifiedin decreasing order of importance: spacebar width, travel, key gap height, key gap width, vertical center-to-center-distance between keys, number of rows, trip force, key curvature and tactile slope.
3. Spacebar width has been identified as the most important physical design variables affecting performance as well as users' satisfaction (from both quantitative and qualitative analyses).
4. Component t1: high negative correlations of gap width and gap height and strong positive correlation of curvature, spacebar width, tactile slope and trip force with t1. The dependent variables speed and accuracy also correlated highly with t1.
5. Component t2: The correlation of accuracy was negative while the other dependent variables (speed, ease of typing and typing satisfaction) were positive. Thus, the linear variate of the independent variables affected accuracy in an opposite direction than the speed, ease of typing and typing satisfaction. Travel, Trip force, height, tactile slope, letter key height and curvature varied negatively with speed, ease of typing and satisfaction.

Table 2. Standardized coefficients of physical variables from PLS Regression (2 components)

Physical Variables	Speed	Accuracy	Ease	Satisfaction	Rank
FMax(grams)	-0.027	0.101	-0.122	-0.107	7
Tallness(.001)					
Travel(.001)	-0.165	0.018	-0.282	-0.258	2
TactileSlope	0.046	0.142	-0.035	-0.025	9
TactileRatio			-0.155	-0.141	
Letter Key Width			0.141	0.13	
Letter Key Height					
CCD Key horizontal					
CCD Key vertical	-0.153	-0.11	-0.163	-0.155	5
Keyboard Width					
Keyboard Height					
Spacebar Width	0.266	0.16	0.308	0.29	1
Gap Width	-0.193	-0.164	-0.186	-0.178	4
Gap Height	-0.206	-0.174	-0.199	-0.191	3
Curvature	0.098	0.148	0.045	0.048	8
# of keys					
# of rows	0.131	0.098	0.138	0.131	6

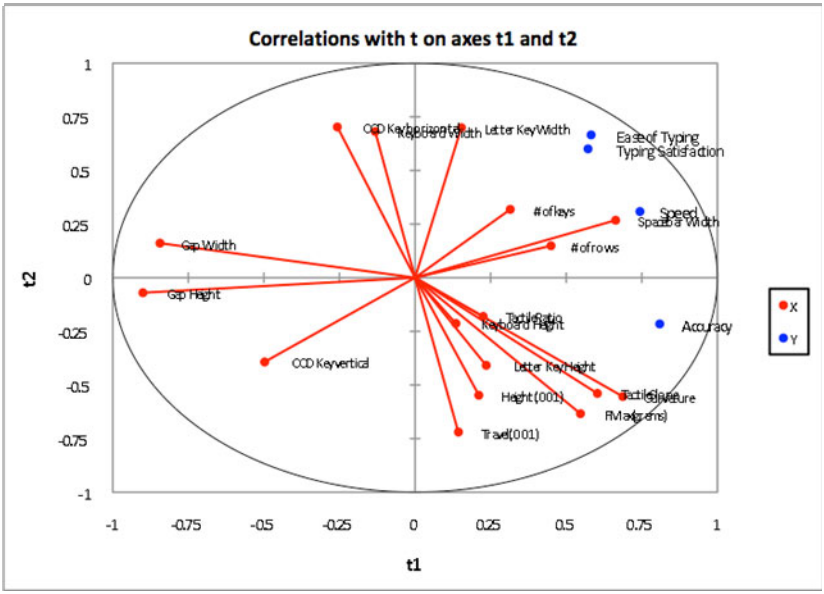


Fig. 3. Correlations of dependent and independent variables with t on axes t1 and t2

3 Mobile QWERTY Keyboard Layout Research

For mobile QWERTY keyboard layout research, the layouts can be constructed from either paper/foam models or from a QWERTY keyboard construction tool to be used in hand-held Windows XP and Android devices. The use of these tools allows for systematic variations and control of different design parameters and layouts to minimize confounding results. Ultra Mobile PCs with 4" touch display have been used to study different layouts keeping the form factor of the testing devices constant.

Fig. 4. Example of QWERTY layout on an ultra mobile PC with touch display

Example research topics include keyboard size [4], number of rows and columns of QWERTY keys, symbols and numbers placement and mapping, staggered and grid layout. The intent of using virtual QWERTY keyboards in these research studies is to isolate the layout of the keyboard from the effects of the physical, mechanical construction of the keys (e.g., trip force, key height and 3-D contour).

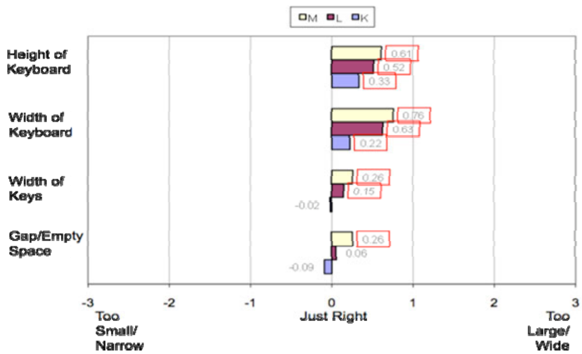


Fig. 5. Example findings of key design assessment. Layouts M, L and K correspond to QWERTY keyboards of decreasing size.

Findings from these research studies suggest that users prefer a separate number row like PC keyboard and 5-row mobile QWERTY keyboards are generally more design "fault-tolerant". Respondents also performed better in grid layout than staggered layout.

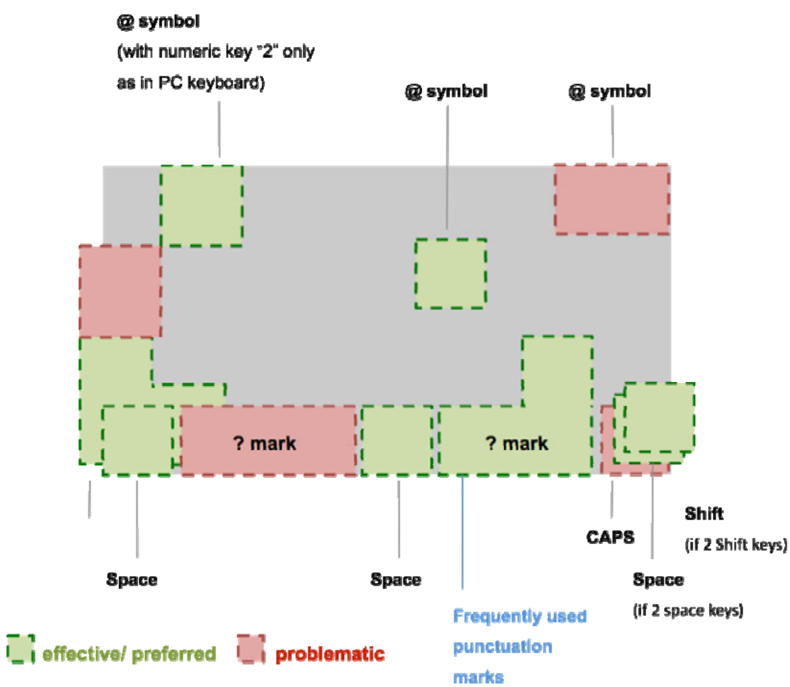


Fig. 6. Example findings of key placement locations in a typical QWERTY layout study

We also used this approach extensively to benchmark QWERTY performance and satisfaction scores of different QWERTY layouts from competitive devices as well as from the requirements of different carriers.

4 Virtual Mobile QWERTY Layout Research

User performance and preference of mobile virtual QWERTY keyboards with different display sizes (and/or varying from factors) were investigated consistently by using the same set of smartphones/tablets (with different display sizes). In order to isolate the impact of differences in touchscreen performance, general user interface as well as the influence of branding, the Keyboard layouts from Apple iphone and iPad were prepared with Samsung virtual keyboard layout designs (for Galaxy S smartphones and tablets). Human performance and satisfaction with typing were captured by having respondents (N=48) enter mixed text, number and symbols sentences in devices of different size display [5].

The virtual keyboard layouts of Samsung Galaxy S smartphones and tablets typically outperformed Apple iphone 4 and iPad keyboard layouts in both satisfaction and ease of use ratings as well as ranking except in typing speed of “words” (letters, numbers and period). Even in these circumstances, the error rate of Apple iphone4 and iPad keyboard layouts was typically larger than Samsung keyboard layouts.

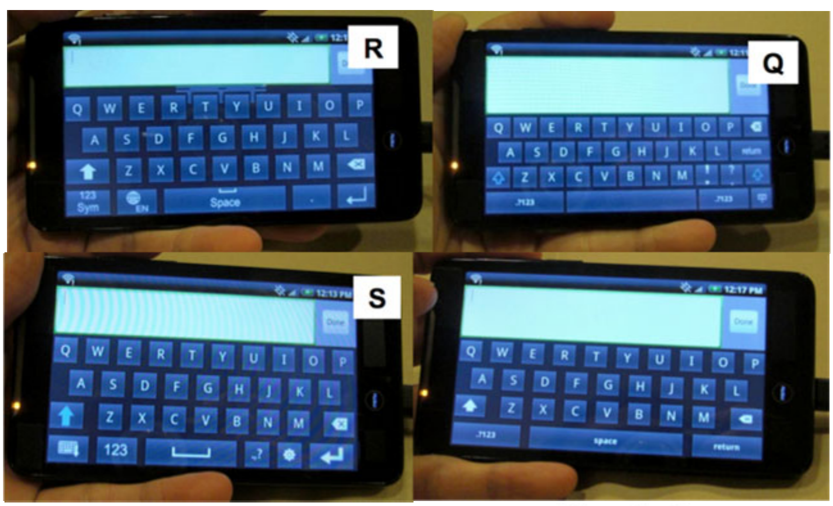


Fig. 7. Examples of different keyboard layout on a 4.3” display smartphone

User satisfaction for portrait virtual keyboard layouts increased significantly or directionally with increasing touch display size (significant results were observed in all the keyboards tested). The effect was much less pronounced with landscape virtual keyboard layouts (significant results only observed in Samsung Tablet).

Respondents made significantly or marginally less errors using Samsung Smartphone and/or Tablet keyboard layouts compared to Apple iPad and/or iphone keyboards.

Display size as a factor is also significant in the number of errors of both Samsung Smartphone and tablet Portrait layouts from 4” to 4.3” display. These results pointed towards the use of 4.3” display for portrait QWERTY layout would significantly reduce the number of typing errors compared to 4.0” display in typical text typing task (but not in mixed entry).

5 Future Direction

We will continue to benchmark performance and user satisfaction of mobile QWERTY keyboards when new devices are launched into the market and also test new keyboard layouts and input paradigms through the techniques we covered above. One possible direction for future research is to investigate the use of new predictive or adaptive text input methods on top of the QWERTY keyboards.

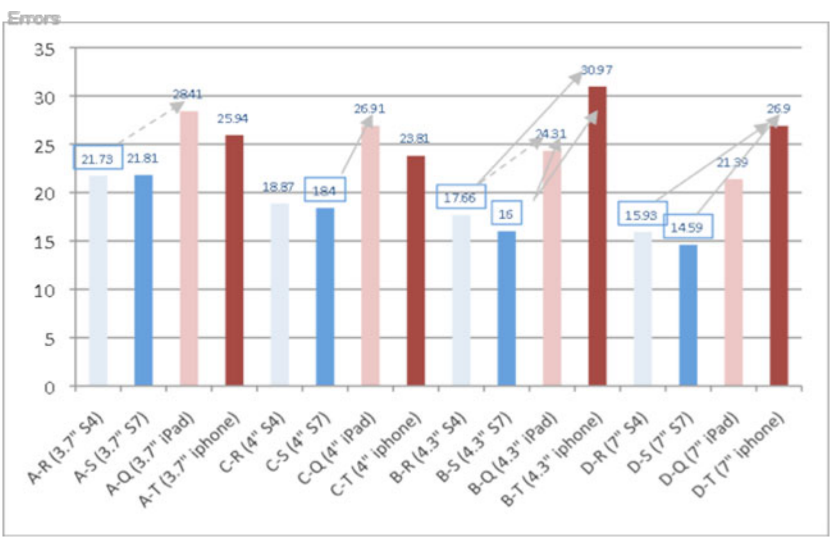


Fig. 8. Typing errors of Samsung keyboards significantly lower than Apple keyboards (Landscape mode, text/symbols/numbers entry). S4 Samsung Galaxy Handset, S7 Samsung Galaxy Tab.

Acknowledgments. The research studies were conducted and analyzed either by the user research staff from STA or through third-party vendors – HumanCentric, Gist Design or Strategy Analytics.

References

1. MacKenzie, I.S., Tanaka-Ishii, K.: Text Entry Systems: Mobility, Accessibility, Universality. Morgan Kaufmann, San Francisco (2007)
2. Hung, Y.S., Vance, S.: QWERTY Keyboard Layout Evaluation (LG Ally, HTC H2 and Samsung EPIC). User Research Report, Samsung Telecommunication America (2010)
3. Hung, Y.S., Vance, S., Kang, N.E.: Mobile QWERTY Keyboard Best Practice: Final Research Report & Physical Design Recommendations. Samsung Telecommunication America & Samsung Electronics Ltd (2009)
4. Hung, Y.S., Vance, S.: QWERTY Keyboard Layout Evaluation (Samsung 5-line Keyboards in different dimensions). User Research Report, Samsung Telecommunication America (2010)
5. Hung, Y.S., Yang, S., Vance, S.: Virtual QWERTY Keyboard Layout Evaluation (Samsung Galaxy Handsets & Tablets, Apple iphone and iPad). User Research Report, Samsung Telecommunication America (2010)

Comparisons on *Kansei* Factors of Attractiveness between Initial and Long Term Use of Mobile Phone

Nanae Kashiwase[1], Keiko Kasamatsu[1], Hiroaki Kiso[2]
Shin'ichi Fukuzumi[2], Misako Yamagishi[3], and Hideo Jingu[3]

[1] Tokyo Metropolitan University,
6-6, Asahigaoka, Hino-city, Tokyo, 191-0065, Japan
kashiwase_nanae@yahoo.co.jp, kasamatu@sd.tmu.ac.jp
[2] Information and Media Processing Laboratories, NEC Corporation
1753, Shimonumabe, Nakaharaku, Kawasaki-city, Kanagawa, 211-8666, Japan
h-kiso@ah.jp.nec.com, s-fukuzumi@aj.jp.nec.com
[3] Kanazawa Institute of Technology,
3-1, Yatsukaho, Hakusan-city, Ishikawa, 924-0838, Japan
{yamagishi-m,jinguh}@neptune.kanazawa-it.ac.jp

Abstract. Many information devices have high performance and functionality by the progress of information and technology and have made lives of people more convenient. Although high functionality is a cause of the attractiveness, it is difficult to satisfy the user only by this. Therefore, the purpose on this research was the comparisons on *Kansei* factors of attractiveness between initial and long term use of mobile phone. As the results, the four factors which are operability factor, motivation factor, familiarity factor, and accuracy factor were extracted as elements of the attractiveness on initial use. Moreover, the two factors, activeness factor and affinity factor were extracted as elements of the attractiveness on long term use. Therefore, it was suggested that the criteria which evaluate the products change as follows. It evaluates to the detail on initial use, and it evaluates it through the whole on long term use.

Keywords: *kansei*, attractiveness, usability, mobile phone, use term.

1 Introduction

The product or system designers have passion which they want to make good products. The passion is a driving force for the designers to produce products. The Japanese transcendent technologies on manufacturing have been boosted up by the passion and the Japanese products have been become high functionality and multiple functions. At the same time, it has responded to the feeling of the user who wants to use high-performance and multifunctional products, and it was considered to be a good product. However, while a competitive battle for market shares was repeated, products which did not have high efficiency and various functions disappeared from the market. Currently the user has enjoyed high-performance and multifunctional products, it is ordinary for users in the market where the change is intense to the improved performance.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 593–601, 2011.
© Springer-Verlag Berlin Heidelberg 2011

The Japanese making products newly are aiming to investigating attractiveness which emotionally engaging in addition to high-performance and multifunctional products. Norman, D.A[1),2)] has discussed the importance of appealing to emotion like pleasure and joy on a product. Furthermore, not only the function but the degree of satisfaction which turned the viewpoint to a user's emotion was included in the concept of the usability[3),4)] on the product use.

Therefore, the attractiveness that a user feels on product use was examined in this research.

2 Purpose

The purpose on this research was the comparisons on *Kansei* factors of attractiveness between initial and long term use of mobile phone. The target term of this research was limited at the time of use of a product. It is possible to offer better product to user by incorporating elements which user feels as attractiveness at using a product. Moreover, it is possible to produce products which continue to be supported positively by user.

3 Methods

We hypothesized that attractiveness changed by period of the use. Thus, the use phase of the mobile phone was divided in two periods of initial and the long term use. The period of the use in initial use was defined as the phase of purchase examination in this research. The period of the use in long term use was defined as one year or more. This research was performed experiment and corresponding to each use phase.

4 The Method of Experiment Ⅰ

4.1 Purpose of Experiment Ⅰ

The experiment was aimed to specify the elements of the attractiveness to feel through product use on initial use for user.

4.2 Participants

A total of twenty seven participants took part in the experiment . Twelve men were 20 23 years old. Fifteen women were 20 22 years old. Since initial users were the object of experiment , it limited to users who had not used the mobile phone used on the experiment.

4.3 Experimental Samples

The samples were three models of mobile phones. The characteristics of three mobile phone models were as follows; Model ① was touch panel of whole surface. It could operate by SW key of all except a hardware key to return home. Model could be operated by the HW keys, such as a ten key, like many existing terminals, and has the feature in how to open a display. Model had touch panel on a whole surface, but an HW key was filled up in comparison with model ① (ON/OFF hook, menu etc.) and had a characteristic pointing device.

4.4 Experimental Tasks

The experimental tasks were five tasks with high frequency in daily life for three models. Five tasks were new creation and storage of e-mail, dialing, setting alarms, taking and saving pictures, and calculation with a calculator.

The participants evaluated based on those operations. The participants performed tasks of experiment by the setup "The user tries to use on the store front to renew the mobile phone on future" or "try to use a friend's mobile phone".

4.5 *Kansei* Words in Questionnaire

The evaluation items were selected 303 words refer to Saito et al.[5)], and Tagawa[6)] as word related to usability on mobile phone. Next, the similar items and items with weak related to usability were eliminated by brainstorming. The subjective evaluation items were composed of 107 evaluation words as *Kansei* items and 4 evaluation words as attractiveness items.

5 Results on Experiment Ⅰ

Factor analysis and multiple regression analysis were performed on experiment .

5.1 The Result of Factor Analysis

The factor analysis was examined to explore elements of attractiveness on experiment and integrate the variables for 107 items except 4 items about attractiveness. At first, the number of items was selected by exploratory factor analysis using principal factor analysis. After that, the factor was extracted by confirmatory factor analysis. The result of cumulative contribution ratio on confirmatory factor analysis (principal factor analysis, varimax rotation) was 70.53%. Four factors were extracted as the follows (Table 1);

Table 1. The factor analysis result of experiment

	Factor 1 Operability	Factor 2 Motivation	Factor 3 Familiarity	Factor 4 Accuracy
easy-to-understand	.869	.031	.205	.026
practical	.837	.123	.164	.030
control at will	.831	-.034	.213	.067
without a sense of discomfort	.792	.028	.095	-.111
comfortable	.783	.299	.229	-.077
convenient	.783	.392	.149	.164
good operability	.768	.413	.179	.038
Learnability	.763	.308	.119	.130
not irritability	.746	.263	.126	.200
considerate	.719	.274	.095	.065
hard to make an error	.718	-.130	.036	.251
considerate	.717	.417	.116	-.095
a sense of safety	.713	-.085	.342	.411
easy to find	.711	.051	.033	.046
efficient	.695	.205	.124	.229
useful	.683	.410	-.194	.055
lean	.646	.190	-.175	-.047
satisfaction	.621	.610	.036	.048
delightful	.590	.531	.184	-.093
favorable impression	.588	.562	.311	-.050
simple	.583	.152	.050	.129
clear	.501	-.008	-.123	.158
feel attachment	.501	.479	.320	-.068
interesting	.357	.803	.190	-.015
fashion	.042	.801	.039	-.040
fresh	-.068	.797	-.059	.012
exciting	.251	.797	.156	-.064
originality	.029	.794	-.074	.126
affecting	.163	.787	-.128	.187
well-designed	.295	.785	-.138	.024
want to touch	.434	.763	-.012	.037
unusual	-.230	.746	-.139	.078
pleasant	.444	.696	.190	-.206
free	.211	.692	.089	-.076
refreshing	.529	.680	.229	-.126
excellent	.456	.674	.024	.295
kinetic	-.263	.663	.140	.107
beautiful	.367	.641	-.053	-.097
high-performance	.434	.632	-.195	.078
clean	.327	.605	.069	.059
high quality	.016	.598	.016	.426
intelligence	.436	.524	-.234	.408
intuitive	.106	.521	.118	-.134
bright	.059	.506	.123	.111
standard	.472	-.486	.319	.330
friendly	.488	.123	.714	.033
casual	.435	.146	.659	-.084
accurate	.469	-.084	.170	.709
grownup	.155	.219	-.409	.624
cumulative contribution ratio（%）	29.50	55.79	60.59	64.89

Factor 1: Operability

There were evaluation items for operation, such as “control at will”, “easy to use”, and “comfortable” etc. on operability factor.

Factor 2: Motivation

There were evaluation items for the movement of mind related to operation, such as “interesting”, “fresh”, and “exciting” etc. on motivation factor.

Factor 3: Familiarity

There were evaluation items for the approachability, such as “friendly” and “casual” on familiarity factor.

Factor 4: Accuracy

There were evaluation items for the accuracy, such as “accurate” and “grownup” on accuracy factor.

5.2 The Results of Multiple Regression Analysis

The factors were examined to prescribe the attractiveness by multiple regression analysis using stepwise way. The objective variables were “attractive”, “want to buy”, “want”, and “want to use” as attractiveness items. The explanatory variable was each operability factor, motivation factor, familiarity factor, and accuracy factor. As the results, the attractiveness was composed of operability factor and motivation factor.

-Multiple regression equations-

“Attractive”=0.832×”operability factor” 1.040×”motivation factor”+4.148 (R^2=0.819)
“want” 0.852×”operability factor” +1.220×”motivation factor”+3.741 (R^2=0.766)
“want to use” 0.926×”operability factor” +1.240×”motivation factor”+3.889 (R^2=0.724)
“want to buy” 0.753×”operability factor”+1.001×”motivation factor”+3.333 (R^2=0.600)

6 The Method of Experiment Ⅱ

6.1 Purpose of Experiment

The experiment was aimed to specify the elements of the attractiveness which feel through use of the product during long term for user. In addition, the elements of the attractiveness of the experiment were compared with that of experiment .

6.2 Participants

A total of twenty participants took part in the experiment . Ten men were 20 24 years old. Ten women were 21 23 years old. Since the object of experiment was long term users, it limited to users who had used the mobile phone owning now more than one year.

6.3 Experimental Samples

The samples on the experiment were participants' mobile phones. No limits were set for telecom companies or models.

6.4 Experimental Tasks

The participants evaluated same five kinds of tasks on the experiment .

6.5 *Kansei* Words in Questionnaire

The *Kansei* words of experiment were 112 words that added "want to keep using" to 111 words of experiment .

7 Results on Experiment Ⅱ

Factor analysis and multiple regression analysis were performed on experiment .

7.1 The Result of Factor Analysis

The factor analysis was examined to explore elements of attractiveness on experiment and integrate the variables for 107 items except 5 items about attractiveness. At first, the number of items were selected by exploratory factor analysis using principal factor analysis. After that, the factor was extracted by confirmatory factor analysis. The result of cumulative contribution ratio on confirmatory factor analysis (principal factor analysis, varimax rotation) was 75.68%. Four factors were extracted as the follows (Table 2);

Table 2. The factor analysis result of experiment Ⅱ

	Factor 1 Activeness	Factor 2 Affinity
surprise	.969	.047
impressive	.925	-.006
stereoscopic	.882	.184
smooth	.837	.245
yearning	.758	.270
motivated	.632	.615
without a sense of discomfort	-.111	.891
casual	.312	.882
stress-free	.001	.855
feel good	.413	.820
free	.177	.652
natural	.487	.581
cumulative contribution ratio (%)	39.96	75.68

Factor 1: Activeness

There were evaluation items which the operation appealed positively to the emotion, such as “surprise”, ” impressive”, and “three dimensional” etc. on activeness factor.

Factor 2: Affinity

There were evaluation items for stress-free to operation, such as “without a sense of discomfort”, “casual”, and “stress-free” etc. on affinity factor.

7.2 Result of Multiple Regression Analysis

The factors were examined to prescribe the attractiveness by multiple regression analysis using stepwise way. The objective variables were “attractive”, “want to buy”, “want”, and “want to use” in addition to “want to keep using” as attractiveness items. The explanatory variables were activeness factor and affinity factor.

-Multiple and single regression equations-

“Attractive”=0.638×“activeness factor” 0.546×“affinity factor”+4.200 (R^2=0.519)
“want” 0.661×“activeness factor” +4.100 (R^2=0.377)
“want to use” 0.741×“activeness factor” +0.376×“affinity factor”+4.350 (R^2=0.754)
“want to continue using” 0.702×“affinity factor” +4.200 (R^2=0.274)

The regression expression on “want to buy” wasn’t obtained.

8 Discussion

8.1 Factor Analysis

As the result of both Experiment and , it was recognized there were the factors about user-friendliness and comfortable, and about the desire and interest for operation. First of all, the factors about user-friendliness and comfortable for initial use and long term use were examined. For initial use, the items which were evaluating operation like “easy-to-understand”, “practical” were obtained as the operability factor. In comparison, for long term use, there were the items how the user felt the operation, rather than the evaluation for the operation such as “without a sense of discomfort”, “casual” and “stress-free”. In addition, there were evaluation items for detail of operation, such as “learnability” and “hard to make an error” etc. in initial use, however not long term use. These showed that there was the change of sense on operation from details of product to entire impression by use phase.

Next, the factor about the pleasure and interest of operation for initial use and long term use were examined. For initial use, the items which were evaluated an impression concerning the operation known by intuition was expressed in detail like “interesting”, “fresh” and “exciting” were obtained as the motivation factor. On the other hand, there were evaluation items for whole operation, such as “surprise”, and “impressive” etc. on activeness factor of long term use. Therefore, it recognized that there were differences for sharpness of impression on operation and collective impression between initial use and long term use. The evaluation to the detail of operability didn't appear in long term use. That is to say, it was thought that the user did not worry about the detail of the operation because the user became accustomed to the detail operation through use.

8.2 Multiple Regression Analysis

As the results of experiment , the coefficient on motivation factor was higher than operability factor on the attractiveness items; “attractive”, “want”, “want to use”, and “want to buy”. It was suggested that user attached importance to the emotion like interesting and stylish etc. rather than the operation like easy to understand and practical etc. for attractiveness in initial use. As the results of experiment , the coefficient of activeness factor on “attractive” and “want to use” was higher than that of affinity factor. And “want” was composed with activeness factor. On the other hand, “want to continue using” had been composed with only affinity factor. It is necessary to decide the directionality of the operation added to the product according to what desire you would like to hold when the product is used for the user.

9 Conclusion

The purpose of this research was comparing the *Kansei* factors of attractiveness between initial and long term use on mobile phone. The four factors, operability factor, motivation factor, familiarity factor, and accuracy factor were drawn from the result of experiment as elements of the attractiveness on initial use. The two factors, activeness factor and affinity factor were drawn from the result of experiment as elements of the attractiveness on long term use. Therefore, it was suggested that the criteria which evaluate the products change as follows. It evaluates to the detail on initial use, and it evaluates it through the whole on long term use. In addition, the factors of desire and exciting have stronger association with the evaluation of attractiveness than the factors of sense of safety and feeling of relaxation both in initial and long term use. It is necessary to expand the kind of equipments in order to apply to various products, although the target in this research was experimented for the mobile phone. Furthermore, it is necessary to extend age group, diversity of culture.

As also in the introduction, Japan has led the world by superior technology to create products. The point that Japan aims at next will be "technology full of hearts" called heart of the hospitality for a long time in Japan. We hope that the Japanese's delicate sensitivity is made use of, and touched of user's detailed mind is caught and it builds it in the product. We would like to contribute to improvement in technology full of hearts.

References

1. Norman, D.A.: Emotional design: Why We Love (or Hate) Everyday Things, pp. 101–115. Basic Books, New York (2004)
2. Norman, D.A.: The Transmedia Design Challenge: Technology that is Pleasurable and Satisfying. Interactions 17(1), 12–15 (2010)

3. ISO9241-11, Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability (1998)
4. Nielsen, J.: Usability Engineering, pp. 23–48. Morgan Kaufmann, San Francisco (1993)
5. Saito, E., Tagawa, T.: Selection of Sensory Evaluation Terms on Design, and Effectiveness of Those Terms by Sensory Evaluation on Design of Article. Japanese Journal of Sensory Evaluation 3(2), 105–114 (1999) (in Japanese)
6. Tagawa, T., Taguchi, M., Koyama, K.: On the Structure of Semantic Space for Evaluating Commercial Goods. Japanese Journal of Sensory Evaluation 3(2), 115–120 (1999) (in Japanese)

Clarification of *Kansei* Elements of Attractiveness Related to Usability for Long Term Mobile Phone Users

Hiroaki Kiso[1], Nanae Kashiwase[2], Keiko Kasamatsu[2], Misako Yamagishi[3], Hideo Jingu[3], and Shin'ichi Fukuzumi[1]

[1] Information and Media Processing Laboratories, NEC Corporation
1753, Shimonumabe, Nakahara-ku, Kawasaki-shi, Kanagawa, Japan
[2] Tokyo Metropolitan University
6-6, Asahigaoka, Hino-shi, Tokyo, Japan
[3] Kanazawa Institute of Technology
3-1, Yatsukaho, Hakusan-shi, Ishikawa, Japan
h-kiso@ah.jp.nec.com, s-fukuzumi@aj.jp.nec.com,
kashiwase_nanae@yahoo.co.jp,
kasamatu@sd.tmu.ac.jp,
{yamagishi-m,jinguh}@neptune.kanazawa-it.ac.jp

Abstract. Recently, in the IT products area, it has become difficult to distinguish among products by function. Usability is one of the most important user requirements for products. And we assume that the other important factor related to distinctiveness among products is *Kansei* (human sensibility). One of the elements for usability from the viewpoint of *Kansei* is attractiveness. IT product manufacturers need to consider the user's *Kansei* when developing products. By clarifying the mechanism involved in *Kansei* and by developing a quantitative *Kansei* element evaluation method, we will be able to develop attractive IT products continuously. We targeted long term users, and we ran a subjective evaluation experiment for attractiveness related to usability including a questionnaire concerning the factors that influenced the judgment of the evaluation and we clarified some of the elements of human interface that influenced *Kansei* elements of attractiveness related to usability. The results indicate that when people use mobile phones considering accessibility to some function, they feel "kinetic", "dynamic", and "refreshing" and when they use mobile phones considering errors when sending an e-mail or making a call, they feel "three dimensional", "impressive", "surprising", "epoch-making", and "natural" and want to use them. The results also indicate that when they use mobile phones considering ease of inputting text, they feel "good operability", "hard to make an error", "controllable as one likes", "stress-free", and "want to keep using".

Keywords: *Kansei*, attractiveness, usability, mobile phone, long term use

1 Introduction

Recently, in the IT products area, it has become difficult to distinguish among products by function. Usability is one of the most important user requirements for

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 602–610, 2011.

© Springer-Verlag Berlin Heidelberg 2011

products. And we assume that the other important factor related to distinctiveness among products is *Kansei* (human sensibility). One of the elements for usability from the viewpoint of *Kansei* is attractiveness. IT product manufacturers need to consider the user's *Kansei* when developing products. To make attractive products, it is necessary to determine how to evaluate attractiveness quantitatively. By clarifying the mechanism involved in *Kansei* and by developing a quantitative *Kansei* element evaluation method, we will be able to develop attractive IT products continuously. In a previous study[1)], the authors analyzed an evaluation model for long term users who use interactive products and a satisfaction evaluation structure. However the problem remains that we can't evaluate the elements of the satisfaction evaluation structure quantitatively. In other previous studies[2),3),4),5),6),7)], some developed new methods to understand quantitatively a variety of *Kansei* words, some analyzed if they can apply physical elements to *Kansei* words, some proposed an analysis technology for a psychological evaluation structure, and some discussed a relationship between subjective usability and objective usability. But their results have not solved issues of how to evaluate attractiveness quantitatively.

To clarify these issues, we studied the following approach.

1. We dealt with two user segments. One was initial users. The other was long term users. The reason that we used these segments was that depending on the period of having the products, the points of evaluation for attractiveness related to usability may be different.
2. In the study described in this paper, we targeted long term use, and we ran a subjective evaluation experiment for attractiveness related to usability including a questionnaire concerning the factors that influenced the judgment of the evaluation. On the one hand, we made a factor analysis of the results of the subjective evaluation experiment, and we clarified *Kansei* factors of attractiveness related to usability. On the other hand, we clarified some of the elements of human interface influenced *Kansei* elements of attractiveness related to usability.
3. We paid attention to physiological data measurement as a quantitative *Kansei* element evaluation method, and analyzed the correlation of the results of a subjective evaluation experiment for attractiveness related to usability and physiological data measurement. If there is a correlation between the results of subjective evaluation experiments and physiological data, we can evaluate attractiveness related to usability by using physiological data.

2 Method

A total of 20 participants took part in the experiment. The participants were university students, long term users who have used their mobile phones for one year or more. The participants tried five common tasks (to make an e-mail and save it, to make a call, to set an alarm, to take a photo and save it, and to calculate with a calculator) with their mobile phones, and after that they filled in a subjective evaluation and a questionnaire. They were directed to evaluate on the basis of neither price nor brand, but usability. The subjective evaluation terms were composed of 112 evaluation words which we picked out in the preceding study[8),9),10)]. They filled in a questionnaire and gave the reasons why they "felt easy to use", "felt difficult to use", "wanted to keep using", and "didn't want to keep using", for each task.

3 Results

From the results of the questionnaire concerning the question about "want to keep using", we focused on accessibility to some functions concerning e-mail, phone, alarm, camera, and calculator. And we found that mobile phones were divided into 2 groups. One group (1) was composed of the phones they wanted to keep using by accessibility to the function in each task. The other group (2) was composed of the phones they did not want to keep using by accessibility to the function in each task. Group (1) and (2) are not the same number of people (unbalanced design).

Table 1. Result of the questionnaire -"want to keep using" because of ease of access to some functions

c																				
c c																				

From the results of the questionnaire concerning the question about "don't want to keep using", we focused on errors (when sending an e-mail or making a call). And we found that mobile phones were divided into 2 groups. One group (3) was composed of the phones they didn't want to keep using because of errors. The other group (4) was composed of the phones they wanted to keep using or didn't want to keep using because of anything other than errors. Group (3) and (4) are not the same number of people (unbalanced design).

Table 2. Result of the questionnaire -"don't want to keep using" because of errors

From the results of the questionnaire concerning the question about "feel easy to use" and "feel difficult to use", we focused on ease of inputting text. One group (5) was composed of the phones they felt easy to use due to being easy to inputting text. The other group (6) was composed of the phones they felt difficult to use due to not being easy to inputting text. Group (5) and (6) are not the same number of people (unbalanced design).

Table 3. Result of the questionnaire -"feel easy to use" and "feel difficult to use" because of ease of inputting text

		B				F		H		K										
y																				
y																				
g	)	5)	5)	5)	)	)	)	5)	)	)	5)	5)	5)	5)				)	5)	)

4 Discussion

As the factor of accessibility to some functions with two levels wherein accessibility to the function in each task made users either want to keep using or not, we conducted one-way ANOVA. As a result, a significant difference was shown at the significance level of 5% for "kinetic", "dynamic", and "refreshing". Therefore, the results indicated that when they use mobile phones, when considering accessibility to some functions, they feel "kinetic", "dynamic", and "refreshing".

Table 4. Result of the subjective evaluation -kinetic, dynamic, refreshing

s													
	u (1						u (						
					N								
k	7			4				4		4			
dy				4					1				
sh				4	4		4					4	

(1 7 s l s

Table 5. Result of the ANOVA –kinetic

Table 6. Result of the ANOVA -dynamic-

y m				

Table 7. Result of the ANOVA -refreshing-

	SS		S	
r r s n	5 5	1	5 5	5 1
rr r	1 1	1 1	1	
t t l	1	1		

p 5

The reason why they feel "kinetic" and "dynamic" is that they feel the expression spoken of as "accessibility to some functions" to be movement and they gave high points to "kinetic" and "dynamic" as expressions concerning movement in 112 evaluation words.

As the factor of error operation with two levels wherein errors when sending an e-mail or making a call didn't make users neither want to keep using nor not, we conducted one-way ANOVA. As a result, a significant difference was shown at the significance level of 1% for "three dimensional", "impressive", "surprising", "epoch-making", "want to use", and "natural". Therefore, the results indicated that when they use mobile phones, when considering errors when sending an e-mail or when making a call, they feel "three dimensional", "impressive", "surprising", "epoch-making", and "natural" and want to use.

Table 8. Result of the subjective evalution -three dimensional, impressive, surprising, epoch-making, want to use, natural-

s r	t r																			
	r						r													
		H						C	D		F	G		J	K		N	P		S
t r s			4	4			4	4					4							1
r ss v	7	4		4					4			4		1	4	4				
s r rs			4	4	4								4	1						1
	7		4	4			4						4	1						4
tt s										4		4	4		4	4	4		4	
t r								4		4		4	4	4	4	4			4	4

r t s s

Table 9. Result of the ANOVA -three dimensional-

Table 10. Result of the ANOVA -impressive-

Table 11. Result of the ANOVA -surprising-

Table 12. Result of the ANOVA -epoch-making-

Table 13. Result of the ANOVA -want to use-

Table 14. Result of the ANOVA -natural-

	SS	d	S	
a ra	1	1	1	1 1 *
rr r	1	1		
a		1		

*p 1

According to the other study in this session (Submission ID: 1158), the result of factor analysis for long term users shows that there are two main factors. One of two main factors is named activeness factor including "three dimensional", "impressive", and "surprising". The other is named affinity factor. This result and the result of ANOVA indicates that there are long term users that don't want to keep using because of errors that trouble another person, such as sending the wrong e-mail or making the wrong call and they think that activeness is more important than affinity.

As the factor of ease of inputting text with two levels wherein ease of inputting text made users feel either easy to use or difficult to use, we conducted one-way ANOVA. As a result, a significant difference was shown at the significance level of 1% for "good operability", "hard to make an error", "controllable as one likes", "stress-free", and "want to keep using". Therefore, the results indicated that when they use mobile phones, when considering errors when sending an e-mail or when making a call, they feel "good operability", "hard to make an error", "controllable as one likes", "stress-free", and want to keep using.

Table 15. Result of the subjective evaluation -good operability, hard to make an error, controllable as one likes, stress-free, want to keep using-

ka e or	a or																
	gro									gro 6							
			D		K												
goo o era l y			5	5					5								
har o ake a error		5		5	5	5		5	5			5	5				1
o rolla le a o e l ke		5		5	5		5	5	5	5		1	5		5		
re ree			5	5				5									
a o kee g				5	5									5			

1 7 ra g ale

Table 16. Result of the ANOVA -good operability-

″ y″					*
	0		0		
	0				

* 0

Table 17. Result of the ANOVA -hard to make an error-

″ n ″					*
			0		
	0				

* 0

Table 18. Result of the ANOVA -controllable as one likes-

″ n n ″	0		0	0 *
	0 0			

* 0

Table 19. Result of the ANOVA - stress-free

″ ″	0		0	0 *
	0		0	

* 0

Table 20. Result of the ANOVA -want to keep using-

″ o u i g″				
o	0		00	
o				

0

The result of ANOVA shows that the *Kansei* words that are significant, distinctive differences are expressions concerning operability and so ease of inputting text is one of the important factors about operability.

5 Conclusion

We targeted long term users, and we ran a subjective evaluation experiment for attractiveness related to usability including a questionnaire concerning the factors that influenced the judgment of the evaluation. As results of the questionnaire concerning the question about "want to keep using", a significant difference was shown at a significance level of 5% for "kinetic", "dynamic", and "refreshing". Therefore, the results indicate that when they use mobile phones considering accessibility to some function, they feel "kinetic", "dynamic", and "refreshing". As results of the questionnaire concerning the question about "don't want to keep using", a significant difference was shown at a significance level of 1% for "three dimensional", "impressive", "surprising", "epoch-making", and "natural" and "want to use". Therefore, the results indicate that when they use mobile phones considering errors when sending an e-mail or when making a call, they feel "three dimensional", "impressive", "surprising", "epoch-making", and "natural" and want to use. As results of the questionnaire concerning the question about "easy to use" or "difficult to use", a significant difference was shown at a significance level of 1% for "good operability", "hard to make an error", "controllable as one likes", "stress-free", and "want to keep using". Therefore, the results indicate that when they use mobile phones considering ease of inputting text, they feel "good operability", "hard to make an error", "controllable as one likes", "stress-free", and "want to keep using".

We clarified *Kansei* elements of attractiveness related to usability and the relationship between some of the elements of human interface and them.

References

1. Ando, M., Kurosu, M.: The User's Evaluation Process Model and the Structure of Satisfaction for Long-term Product Use. Human Interface 9(4), 443–453 (2007) (Japanese)
2. Yanagisawa, H.: Quantification Method of Diverse Kansei Quality. Journal of the Japan Society of Mechanical Engineers 110(1069), 959 (2007) (Japanese)
3. Ando, M.: Self-efficacy and Product Involvement Matrix Analysis (SEPIA). An Attempt at the Analysis of the Product Evaluation as a Result for Long-term Use human interface 11(1), 41–46 (2009) (Japanese)
4. Ujigawa, M.: A consideration of techniques to create attractive spaces and environments: Part1 A proposal Miryoku Engineering, summaries of technical papers of annual meeting architectural institute of Japan. Environmental engineering, 1167–1168 (1994) (Japanese)
5. Tazaki, S., Okazaki, A.: A method to quantify KANSEI directly. In: Proceedings of the 55th annual conference of JSSD, vol. 55, pp. 268–269 (2008) (Japanese)
6. Sauro, J., Lewis, J.R.: Correlations among prototypical usability metrics: evidence for the construct of usability. In: Proc. CHI 2009, pp. 1609–1618. ACM Press, New York (2009)
7. Hornbak, K., Law, E.L.: Meta-Analysis of correlations among usability measures. In: Proc. CHI 2007, pp. 617–626. ACM Press, New York (2007)
8. Kiso, H.: Relations between Kansei Elements of Attraction on Initial Use of Mobile Phone. The Japanese Journal of Ergonomics 46, 360–361 (2010) (Japanese)
9. Kashiwase, N., Kasamatsu, K.: The Consideration for Kansei Factors of Attraction on Initial Use of Mobile Phone. The Japanese Journal of Ergonomics 46, 358–359 (2010) (Japanese)
10. Yamagishi, M., Yammuro, S., Jingu, H.: A Proposal for Indices to Assess Attraction on Initial Use of Mobile Phone. The Japanese Journal of Ergonomics 46, 358–359 (2010) (Japanese)

Suggested Collaborative Learning Conceptual Architecture and Applications for Mobile Devices

Kwang Lee and Abdul Razaque

University of Northern Virginia
7601 Little River Tnpk., Annandale, VA 22003, USA
{Kwang.lee,Arazaque3275}@unva.edu

Abstract. This paper describes the theoretical and technical foundations for designing and developing an effective mobile collaborative learning (MCL) environment. The paper suggests a prototype based on client-server side with the support of functional components and working procedures, which help users to obtain contents from a server to meet the pedagogical needs. The proposed prototype provides the best MCL environment for students who want to learn at home and working places through mobiles. Finally, to find valuable hidden issues, the paper introduces a new application, "group" in Android operating system, and then conducts usability test in order to facilitate users for accessing and obtaining the contents in collaboration learning procedures.

Keywords: Mobile Collaborative Learning (MCL) Environment, Optimization of Architecture, Client & Server side Prototypes for Collaboration.

1 Introduction

The trend of competition has been growing rapidly in global market with advent and deployment of new technologies. The latest technological revolution is the emergence of mobile wireless communication technology. Mobile devices have grown in popularity to become one of the most common consumer devices and ïcheaper hand held device which we can carry and use whole day everywhere. With incorporation of emerging hardware technologies in mobile devices; such as motion sensors, cameras, global positioning system, infra-red, Bluetooth and others supported by broadband connections, mobile devices began to support many different types of education applications. Now they are more flexible to integrate the existing services by employing web based interfaces, so that mobile devices become attractive tools to complete the demand for collaborative learning.

Collaboration has been getting more importance in educational environment and the focus of collaborative learning has been implanted from elementary to higher educational institutions. The concept of mobile-based learning (MCL) is completely different from classroom-based learning method. This pedagogical method of the learning provides many possibilities, such as providing the opportunities to group of people, working in same or different organizations to participate for accomplishment of specific goal using mobile devices.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 611–620, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Therefore, the demand of collaboration learning over the mobile device has been increasing as a major education element. Major research challenges are raised in developing MCL for education such as sharing knowledge, requesting for modified contents, fully accessing to enterprise data warehouse, delivering large rich multimedia contents (video-on-demand), selecting technological components in designing appropriate architecture and adapting application protocols. To support these issues, we propose client and server based prototypes with improving knowledge sharing process, providing access for all users to enterprise data warehouse, containing content-modification facility and delivering large rich multimedia contents. Our contribution will make the application easier for students to obtain succinct information and fast feedback through MCL.

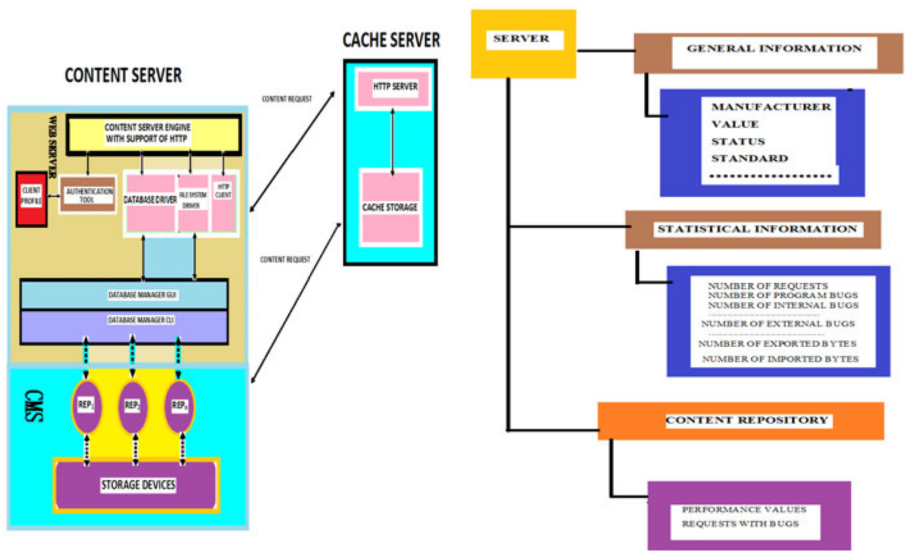
1.1 Background Works

Here, we discuss the salient features of related published work which are contributed to our work. Su, X., and others [1] propose four layer framework prototype for multimedia content generation in mobile collaborative systems. The proposed framework provides the support for users, devices and session management skill. This research offer a novel frame work for multimedia content generation, representation and delivery, for Mobile Collaboration, however the framework is not completely elaborated and lacks incorporation of many components. Zanev, V., Clark, R. [7] suggest the procedures for developing a prototype used for wireless course management system focusing on login and authentication, wireless syllabus, wireless calendar and wireless testing. Also it describes the course contents by explaining how a teacher can interact with students by using HTML interface. Lahner, F. and Nosekabel, H. [8] have implemented the program in University of Regensburg, Germany, which can support e-learning contents to be displayed on computers, and also the system structure provides users with the facility to obtain same contents via mobiles. Marcelo, M. and others [9] have designed C-note in University of Vaxjo in Sweden. Using C-note provides opportunity to collaborate for storing the notes and information in the database on their progressive project. The specialized C-Pen is used to scan the project research material with the support of C-note application.

Barbosa, J. and others [5] proposed the prototype for undergraduate course reference, Grefe, by using mobile and ubiquitous computing. Authors claimed that their approach will improve academic and learning activities. The prototype is based on user profile which stores the information regarding learning process and use location system, so it can identify user physical location and support learning procedures by generic architecture. However, the proposed prototype does not exactly provide the mobile collaborative learning, it suggests generic idea of online learning process. Druin, A. and others [10] have discussed the prototype for their ongoing participatory design project with intergenerational design group to create mobile application and integrate into iP Phone and ipod touch platforms. Bouras, C. and others [11] have introduced INVITE architecture and discussed the user requirements to meet the demand of e-learning in collaborative virtual environment. Even it does not lead to existence of any solid prototype; it can show the technology and standards required for designing MCL.

2 MCL Prototype Structure at Server

To make a successful collaboration system, we need to organize its architecture with latest technologies to meet our expectations. Client-server based prototypes provide smooth transaction of data exchange for collaborative learning environment. These prototypes give the information regarding the course materials, easy access to check the grades and use of labs. Among of various conceptual collaborative architectures have been proposed so for, four layered components of collaborative framework [1] was adapted for the basic structure of this research MCL framework. It consists of four major components, content generation layer, communication layer, content regeneration layer, and content visualization layer in which each layer has been assigned different responsibility. As seen in Fig. 1 (a), at the server side a content server is the main component of content generation layer which is based on content server engine (CSE).



(a) Server side operations (b) Function of server with repository

Fig. 1. Conceptual architecture for MCL

The basic engine of the prototype is implemented on Internet information server; so its function is to receive URLs, checks and triggers (stored procedures) process for the requests. The content server continues to watching the interface whether new request from clients is received or not. If the new request is displayed on web through template, then the rest of team members are informed through email, then the server forwards the requested content to CSE. To verify the status of the client, CSE sends the request to authentication tool to save the profile of legitimate clients and then check the client status. We propose HTTP client and integrate with CSE to save and retrieve data from database driver or file system driver.

Database Manager CLI tool in the server creates monitors, manages backups and restore the instance of database, furthermore the tool supports to interactive and background operations. By using Java, C++, XML and other programming languages, we can access the Database Manager CLI tool by creating the specific programming interfaces, module and documentation which are offered from the internal tool features. Database Manager GUI also supports different features on different types of mobile devices, and also it can basically provide user-friendly interfaces to remote mobile devices for the collaboration easily.

GUI and CLI promote to store and extract the contents forwarded from database or file system drivers by using different repositories. Content repositories receive a content request message from GUI and CLI, and search the requested content in the storage devices of content management system (CMS). The content repositories are also part of the CMS, which supports and functions like logical storage for different storage devices. Also the Repository can be helpful to store different kind of contents in single repository table such as comments, articles, questions, answers, news, tutorials. So, we can switch from Rep_1 to Rep_n in single table for finding the requested contents for the collaboration, and can monitor the performance of each repository with their respective servers as seen in Fig. 1 (b).

3 Proposed Group Application for MCL

MCL is a revolution for education, which allows users to obtain computer-based information through mobile devices. MCL provides various advantages such as context awareness, portability, connectivity and social interaction [2].

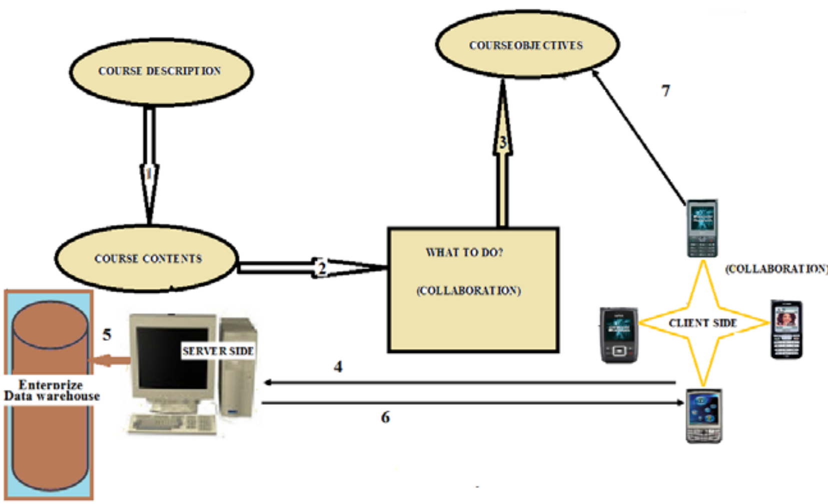


Fig. 2. Process for obtaining the course contents through MCL

Mobile can be a successful tool for the collaboration allowing students to share the information to achieve pedagogical activities. Our prototype for MCL can solve the deep dream of students to make the collaboration with mobile anywhere and anytime.

Our prototype application helps the students to obtain the required contents from enterprise database warehouse (EDW) to meet the course requirements. Suppose, a course “Decision Support System” is offered an online class from the University of Northern Virginia, and some students might not access to their computer during day times under some situation. Instead of using a computer, students can access the server to load their course contents by using mobile devices with MCL environment. Surely to meet the standard of course explained, they can receive class information stored in EDW. The information comprises of Textbook Information, Course Name, Course ID, Course description and others. Hence all these items can be obtained to complete the course by using MCL. Fig. 2 supports the process about the prototype.

To implement MCL prototype on Android mobile operating system which consists of software stack for mobile devices, we incorporate net News Wire software with RSS 2.0 and suggest the following supplementary instructions.

- Add new application with name "group" in application section by using Java programming.
- Resource Manager and Activity Manager should be extended and provided extra responsibilities to control different features of "group" application.
- Libraries section should be modified because it can make easy for structured data storage.
- Display Driver of Linux Kernel to support MCL activities.

Fig. 3. Running group application and calling process on Android operating system

If these instructions are incorporated in Android, a new application "group" with new features will support to obtain the contents from a server to completion of MCL as seen in Fig. 3.

The application "group" is shown with Android OS consisting of control option and delivery option. The control option performs the functionalities of add new contact, delete contact, edit contact, existing collaborative group (C-G) and make new C-G. Delivery option executes the function of receiving and sending information. When clients take the requested contents from a server, they receive through delivery option and saves by using the store component given in Fig. 4.

Fig. 4. Received contents from server are stored in client's mobile

When all of collaborative members of group complete the process of downloading the required contents from RSS feed of a requested client, they starts to follow the process for MCL given in Fig. 5.

Fig. 5. Sharing of file for Mobile Collaborative Learning

If participating members need a file for CL, then they can open the file to receive the knowledge of the contents with reading mode. If they need more information to clear their concept about the topic, then play the videos related information given process.

Fig. 6. Collaboration among the Mobile Collaborative group for achieving the objectives

The group option provides utility to manage the data with its respective options. If data is text based then it should be obtained by using file option, otherwise audio and video options are used. When each of collaborative member receives awareness and obtains knowledge about the topic, subsequently a user start the process of collaboration by using existing collaborative group in Fig. 6.

The process can be supported with H.323 standard published by International Telecommunications Union Telecommunications (ITU-T) to facilitate the real-time audio and video communication. This protocol will help to collaborative members to discuss the contents after reading and watching the video of related information. With integration of "group" application in mobile devices, the course objectives can be met with MCL.

4 Empirical Results to Design Prototype for MCL

One of the highly challenging tasks for designing and developing prototype is to understand how the requirements of users are satisfied by the functionalities of application and architecture. To satisfy the requirements of users, usability testing method shall be conducted based on hybrid heuristic approach. To find meaningful items, we involve 106 users who are including 58 students, 23 teachers, 14 teaching assistants and 11 administrators.

4.1 Testing Conditions

In order to take valuable suggestions for efficient MCL structures, we've surveyed on previous literature and evaluation data [12][13] in first phase, then collect the necessary items through interviews and consult with experts. All items were selected based on the basic requirements of users to meet the pedagogical needs through MCL, then each item was evaluated based on the survey's questionnaire method by the above users.

Table 1. Survey questionnaire for each item

Description of Each Item	**1**	**2**	**3**	**4**	**5**
Should be easy to use	249	0	00	00	00
Asynchronous Collaboration	160	40	30	18	01
Synchronous Collaboration	161	39	29	19	01
To support multimodal MCL	233	13	03	00	00
To provide Archive updating	135	34	61	17	02
Should be User friendly interface	249	00	00	00	00
To give virtual support	139	72	21	10	07

Table 1. (*continued*)

The administrators should be provided the opportunities to record the collaborative activities of students and teachers during the whole session or any specific period of time.	234	06	09	00	00
To provide the opportunities for interactive and shard white board	190	32	22	05	00
Users may need short start time for collaboration	140	30	67	04	08
Server should provide content adoption service	142	28	64	09	06
The Students should have alternative choices in selecting any topic for discussion.	231	09	09	00	00
The Students should have access to check the comments given by teacher regarding their performance and grades.	230	11	08	00	00
The teacher should include critical notes for the performance of each student after completion of MCL session and provide the feedback to improve in future.	232	09	08	00	00
To provide audio and video communication only	192	10	18	19	20
To provide connectivity management support	195	12	32	08	02
To provide the support for session management	193	15	31	07	03
To provide the checking facilities to instructor to check the group members	202	22	14	08	03
To provide the freedom of thoughts to participating group members	211	19	12	07	00
Server should send the message of information updating	184	22	30	10	03
Client should give notification of his/ her availability	156	24	34	28	07
To provide the support for user role	123	43	45	32	06
Portfolio should be created in order to store an information regarding the course	129	56	39	24	01
To include group manager component	124	61	43	21	00
The methods of communication should be direct or mediated	121	60	56	10	02
Instructor should dedicate time to monitor the progress of participating members	201	23	22	03	00
To provide the support to handle the shared information	123	72	54	00	00
To provide privacy and safety	175	32	42	00	00
To provide the facility to contact and invite the participating for collaboration	137	63	49	00	00

Table 1. (*continued*)

The communication should be based on broadcast with support of multicasting	156	67	21	04	01
To make small participating group for collaboration	101	98	41	08	01
To provide support for floor control administration	111	45	52	37	04
To be flexible to collect and extract the data.	247	01	01	00	00
To provide text, graphs, images, audio and video services to meet the requirements of related course of study	245	03	01	00	00
The teachers should have complete access to administer their courses and evaluate the progress of students.	242	04	03	00	00

The way of collecting this basic information is using Face book, Vista survey and personal related group for four days. By using five-point Likert scale, we evaluated each requirement. The five-point Likert items in our questionnaire are specified Strongly Agree=1, Agree=2, Neutral/No Opinion=3, Disagree=4 and Strongly Disagree=5. The results analysis includes answers from all respondents who took your survey in the 3 days period. The items and scale description of is shown in Table 1. Based on the above results, for more testing procedures we will take all of the observations made by the tester, as well as the comments and questionnaire data of the participants, and categorized them. Also we plan to modify and redesign the prototype, and then conduct the final empirical testing on the new prototype again. All these testing procedures will be useful to understand the type of applications required for designing and developing the MCL application.

5 Conclusion and Future Works

The main objective of designing and developing the conceptual-based architectures is to obtain the learning materials on hand-held devices particularly on mobile devices. With applying the prototype architectures, we suggest students to take the contents of the course anywhere and anytime. Here we point out several contributions over the work. First, we have made significant modifications in the four layer architecture previously explained in [7]. Our four layer architecture can provide the efficient and fast way of delivering the contents to mobile node. As this characteristic of four-layered architecture gives fast provision of content to mobile node. Second, we have introduced the client and server based architectural prototype to support to MCL in education. The design and development of client and server based architecture can provide the faster method of delivering the contents to users. Also we have optimized the content server by integrating with a cache server in order to save the time for delivery of data. Third, we have proposed and implemented the prototype application, "group," and recommended some valuable suggestions to use Android operating system and explained the case study how to meet the course objective by using MCL. Finally, we have discussed the user requirements on the basis of survey to design and develop the MCL based architecture to meet the pedagogical requirements of students, teachers, administrative and teacher assistant.

Although this approach is to devise a basic method for efficiently building MCL application, there are still uncovered problems of applying this approach to commercial mobile devices directly. However, our ïarchitecture will properly meet the challenge of MCL and we keep focusing on implementing the whole group application to meet the new needs. We look forward to continuing the research and developing the application according to the future progress of MCL, mobile software and hardware performance. Therefore, the biggest contribution of the paper is to provide a small step for how to design and implement MCL in order to meet ïthe pedagogical needs.

References

1. Su, X., Prabhu, B.S., Chu, C., Gadh, R.: Midleware for Multimedia Mobile Collaborative System. In: 3rd IEEE Wireless Telecommunications Symposium, California (2004)
2. Santos, H.d.F., Santana, L.H.Z., Martins, D.S., De Souza, W.L., do Pardo, A.F., Biajiz, M.: A ubiquitous Computing Environment for Medical Education. In: ACM SAC 2008, Fortaleza, Ceara, Brazil, pp. 1395–1399 (2008)
3. Sung, J.S.: Application Protocol Design for Collaborative Learning. International Journal of u-and e-service science and technology 2(4) (December 2009)
4. Martin, S., Boticki, I., Jacobs, G., Castro, M., Peire, J.: Work in Progress - Support for Mobile Collaborative Learning Applications. In: 40th ASEE/IEEE Frontiers in Education Conference, Washington, DC (2010)
5. Barbosa, J., Hahn, R., Barbosa, D.N.F., Geyer, C.F.R.: Mobile and Ubiquitous Computing in an Innovative Undergraduate Course. In: SIGCSE 2007, Covington, Kentucky, March 7-10 (2007)
6. Jacobs, G.M., Hannah, D.: Combining cooperative learning with reading aloud by teachers. International Journal of English Studies 4, 97–118 (2004)
7. Zanev, V., Clark, R.: 43rd ACM Southeast Conference on Wireless Course Management System, Kennesaw, GA, pp. 2-118–2-123 (2005)
8. Lahner, F., Nosekabel, H.: The Role of Mobile Devices in E-Learning First Experiences with a Wireless E-Learning Environment. In: Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education, Vaxjo, Sweden, August 29-30 (2002)
9. Marcelo, M., Juan, P., Ulrich, H.: C-Notes, Designing a Mobile and Wirelss Application to support Collaborative Knowledge Building. In: Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education, Vaxjo, Sweden, August 29-30 (2002)
10. Druin, A., Benderson, B.B., Quinn, A.: Designing Inter generational Mobile Storytelling. In: IDC 2009-Workshops, Como, Italy, June 3-5 (2009)
11. Bouras, C., Hornig, G., Triantafillou, V., Tsiatsos, T.: Architectures Supporting e-Learning through Collaborative Virtual Environments. In: The Case of INVITE, ICALT 2001, Madison, WI, pp. 13–16 (2001)
12. Hinze-Hoare, V.: Designing a Collaborative Research Environment for Student and their Supervisors (CRESS). School of Electronics and Computer Science, University of Southampton (2007)
13. Ochoa, S., Alarcon, R., Guerrero, L.: Understanding the Relationship between Requirements and Context Elements in Mobile Collaboration. In: Jacko, J.A. (ed.) HCI International 2009. LNCS, vol. 5612, pp. 67–76. Springer, Heidelberg (2009)

Enhancing Interactions of Self-help Groups in Africa through Mobile Solutions: Design Guidelines

Mokeira Masita-Mwangi, Faith Ronoh-Boreh, and Patricia Aruwa

Nokia Research Center Nairobi,
P O Box 29, 00502, Nairobi
mokeira.masita-mwangi@nokia.com,
ext-faith.ronoh-boreh@nokia.com, achiengaruwa@yahoo.com

Abstract. Informal self-help groups (merry-go-rounds) are very popular in Africa. Group members come together for informal banking, social welfare, sharing knowledge, news, ideas and tradition, and income generating projects. In Africa groups tend to be geographically isolated particularly in rural areas. This limits interaction between groups yet this could be beneficial e.g. small group savings put together could create a horizontal microloan system for groups and interest financial institutions to deliver services to them. It would also boost local trade and open new markets for entrepreneurs. Networking self-help groups is thus important and implementation through e-solutions would help overcome geographical dispersion challenges. This paper discusses research findings supporting the case for development and design of e-solutions for linking self-help groups. Further it highlights various opportunities for e-solutions by relevant stakeholders for group linkages and provides some design insights based on design guidelines developed by Nokia Research Center, Africa.

Keywords: Merry-go-round, self-help, group linkages, Africa, social networks, e-solutions, design.

1 Introduction

This paper is based on a recent study of informal self-help groups in Kenya carried out by Nokia Research Centre, Africa. The research was a combination of desk studies and in-depth exploratory and participatory interviews and observations with 10 different groups in different rural parts of Kenya all of which took place between April and May 2009. The study developed an in-depth understanding of the functions, social dynamics and cultures of merry-go-rounds with and went further to demonstrate the extent of use of such informal groups and the potential that exists for networking or collaboration between the groups.

The study results can be deemed to be quite relevant for other African countries despite the wide variety of differences in ethnic groups across Africa. The fundamental roles and needs of these informal groupings remain similar across different geographical regions, cultures and ethnicities in Africa. There are group needs that can be met and

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 621–629, 2011.
© Springer-Verlag Berlin Heidelberg 2011

challenges overcome through sharing or exchange of knowledge, skills, ideas or other form of content between different groups. This paper zeroes in on the importance of group linkages or collaborations as well as the challenges thereof and opens discussion on the possibility of e-solutions in meeting this need and what to take into consideration when designing them.

2 Informal Self-help Groups

Formation of informal self-help groups (also commonly referred to as merry-go-rounds) is one of the main ways through which communities in Africa pool their resources together given that social capital and social ties are important foundations on which communities are grounded in Kenya. Merry-go-rounds are a huge phenomenon in Africa[1]. Self help groups and women groups have the highest numbers amongst the non-profit organizations in Kenya. There were 185,722 self help groups in Kenya and 135,294 women groups by 2005 [1].

There are two main types of informal self-help groups namely Rotating Savings and Credit Associations (ROSCAs) and Accumulating Savings and Credit Associations (ASCAs). The central activity in both of these is savings and loans with the key difference being that in ROSCAs, members save and borrow in a rotational manner from fixed regular contributions of all members while in ASCAs members accumulate a fund from which they can borrow loans at an agreed interest rate. These groups are popular because of the savings flexibility they offer particularly for the poor who cannot afford or easily access formal alternatives. Women who belong to these groups also say that they find it much easier to save as a group as opposed to on their own because of the collective commitment and accountability to the group.

There are also other groups that do not have financial activities; their main function relates to social welfare and other social activities and/or developing and implementing group projects for income generating activities or for self/community development.

3 The Case for Linking Self-help Groups

What some groups have or are very knowledgeable and skilled at, other groups are lacking in [1]. Different groups are at different levels and stages of development and have different experiences. This provides a rich ground for sharing and exchange between the different groups since the fundamental functions of groups are similar. As such members can easily benefit from learning and support across different groups. Development agencies also stand to benefit from successful linkages which they can tap into to deliver various development interventions.

3.1 Resource Pools

Groups can gain access to a larger resource base by pooling their resources. Savings are a good example given that the amounts of money that groups are able to accumulate

[1] Participation is as high as 95% in some countries and in others ranges from 45% to 80% [3].

collectively are indeed substantial[2]. Groups can then borrow funds from this pool. Sharing of other resources such as tools and equipment is also a possibility through established group linkages; many groups are in rural areas and carry out agricultural activities and may therefore exchange or share tools of the trade. What one group has in abundance, another may be lacking in and the reverse is true.

There is also rich knowledge and skills which vary from group to group and the study identified a yearning interest by groups for new ideas and new knowledge. Skills and knowledge in groups are of different categories. There are skills and knowledge held by individual members of a group, collective group skills and community level information. Individuals in groups have various skills such as public speaking, agricultural skills, business management, tailoring, hairdressing, cottage industry skills, counseling housekeeping, midwifery, environmental management and even preservation of the dead.

Existence of an open, increased or enhanced form of linkages and collaborations would enable sharing of skills, knowledge, and groups would more easily implement new activities in their groups. Many self-help group members are micro-entrepreneurs hence the linkages would also create an avenue for them to gain access to new markets for their products.

3.2 Achievement of Community Development Goals

Peer to peer learning could be applied to groups that are linked together basically because individuals in groups would be able to relate quite easily to other individuals like themselves in other groups and appreciate what their peers have achieved in circumstances similar to their own and learn from it. It would quite likely be an inspiring and motivating form of learning. Having realized that self-help groups are a vital channel through which information and knowledge flows in communities, African governments and other development agents have started building capacity in them with a view to enabling them become trainers of other groups. Healthcare and Agriculture are some of the sectors where this is actively on-going. However these development agencies are also limited in capacity and cannot reach as many groups as they would like to hence access to an existing group network would go a long way in increasing their reach.

Through linkage of groups that have attained different progress levels community development needs and goals may be attainable.

3.3 Social/Cultural Integration

Using Kenya as a case example, this African country is largely rural and majority of the informal self-help groups are rural. Kenya has more than 40 ethnic groups which reside in different rural geographical locations within the country. Thus, the members of self-help groups over and above having close social ties also tend to be from the same ethnic community. These ethnic groups speak different languages and have different cultures and practices. Successfully linking or networking the groups would therefore in essence be a positive step towards breaking down stereotypes held by and

---

[2] In 2006 informal groups in Kenya were reported to be mobilizing at least Kshs 1.2 Billion (~ USD 14.2 Million) every month [2].

about different communities and demystifying some cultural practices that may not be widely understood. In addition appreciation of the varied cultural values by different people can only be perceived when groups interact closely and freely with the opportunity to ask questions where they are not sure or clear about why particular things are done in a certain way and not the other. In coming together, social gaps will be bridged and broader networks created. Social gaps exist where communities do not have common understanding on specific issues. In linking groups, ability to bridge the societal differences can be realized with resultant unity across ethnic, cultural, geographical and social boundaries.

4 Challenges in Linking Groups

Varied geographical locations, inherent differences in group structures, functions & attained level of development, socio-cultural differences, as well as patriarchal communities are some of the main challenges identified and concluded in the study that may potentially hinder group linkages.

Currently group networking requires movement from one geographical location to another; the cost element is a bottleneck for majority of groups bearing in mind their socio-economic status. This is therefore very limited at the moment. Also, membership in groups is mostly by women who have a lot of responsibilities on the home front meaning they have limited possibilities for traveling. Men also tend to dominate and take over everything that seems ‘important’ and not allowing the women to participate effectively. As such any networking opportunities would likely be taken over by the men in the case of mixed groups.

In terms of the starting point, what criteria should be used at the onset to determine which groups should be linked to which? How relevant are issues such as number of years in existence, or similarity or difference in group activities in making this decision? It would be important therefore for groups to access sufficient information about other groups in order to decide whether or not to invest their time and other resources in establishing the link and nurturing it. Different groups also have different obligations such as loans, joint projects and investments which may also potentially restrict the networking process. For linkages to work well, these would need to be taken into consideration so that the interests and status of various groups are protected. Collaboration with other groups also has the potential of raising fears associated with loss of group identity or failure to maintain group objectives. The networking platform therefore needs to acknowledge these fears and allow for preservation of individual group identity as well as enable fair sharing of benefits accruing directly as a result of the networks.

Differences in cultural practices across various ethnicities may also be a hindrance to group to group learning. Different values, existing stereotypes, religious beliefs, are just some of the issues that are likely to come up; trying to bring together persons of diverse beliefs may be difficult. Nonetheless motivations for linkages do exist as there are examples of groups comprising mixed ethnicity and religious beliefs. Though this is the exception rather than the norm, it provides hope that there is an opportunity to promote networks across diverse ethnic or religious boundaries.

In order for group networks to effectively meet development needs, the participating groups need to first and foremost be assisted in realizing whatever knowledge and skills they possess and how to leverage these within their own groups and on the networking platform with other groups. Individual group members often do not highly value their skills and knowledge especially that which is acquired informally. They may therefore hold back even when provided with an opportunity to network perhaps believing that they have little or nothing to offer. Sharing and distribution of content is crucial to the success of the network itself since it is likely that users will be seeking to benefit from content they can obtain from as well as the opportunity to share their own. It is therefore important to find a solution for content development on an exchange platform for self-help groups. That which groups possess needs to be harnessed and packaged in a relevant manner taking into account aspects such as language and translation, audio and visual capabilities and the medium through which this content will be distributed. Development agencies also seeking to tap into the networks need to tailor make their content to be easily understood by groups and contextualized to fit their real life situations as well as accepted and utilized.

5 Towards Sustainable Group Linkages

For successful and sustainable implementation of group networks, groups must appreciate the value of the networks from the onset. Sound understanding of how to use the networks for their benefit as groups as well as individual group members is required. These benefits could be financial or social and in the long run could also benefit the community as a whole. Over and above the challenges identified in section four above, the greatest challenge or restrictive factor that would need to be overcome is fear and mistrust associated as well as alienation whereby there is lack of ownership of the networking process by the groups.

5.1 Designing and Developing E-Solutions

There appears to be significant potential for e-solutions in creating a sustainable platform for linking or networking informal self-help groups in Africa. Creating the groups social platform through a web-based service for example would eliminate or mitigate some of the networking challenges discussed such as time and cost elements of wide geographical dispersion and the nature of patriarchal communities. It would also increase transparency and democracy in group affairs since anyone with the desire and interest could access the network without having to break through any physical and bureaucratic barriers associated with a physical networking process which would most likely be controlled by certain influential individuals. An e-solution might also be an easier or softer way to approach the socio-cultural differences since groups may feel less threatened if they can for a start interact with other groups of different cultures in a virtual environment and slowly begin to appreciate and understand their differences. It would also be easier and faster through a web service for instance for groups to identify already which groups exist and by these groups displaying certain information about themselves they would quickly know how similar or different they are from one another and select relevant groups with which they would want to interact.

There are various opportunities in different sectors that can be taken up by the relevant stakeholders and implemented as e-solutions for linking informal self-help groups. One is the web-service as earlier mentioned which would act as a social media platform where groups could input information about their group and search for other groups within the platform. The groups would then be able to communicate virtually and develop some kind of relationship suitable and beneficial to them. Related to this is the opportunity for development of digital content which can be distributed through the social media platform – agriculture, health, general knowledge, adult literacy, business and entrepreneurship, civic education are just some of the topics that would be of interest and developmental value to the groups. Another opportunity would be in the area of service broadcasting. The use of mobile phones and radio programming in interactive format is becoming common place in Kenya. This means participation does not exclude those who are geographically removed from accessing development agencies. As such service broadcasts specifically designed to cater for group needs can be developed for news and information from the groups themselves as well as external development agencies. The possibility of integrating service broadcasts with the social media platform can also be explored. Financial related e-solutions are indeed also worthy of consideration given that financial activity is central to most groups. Once groups begin to interact on a social network, they may wish to start linking together in matters financial. This could be in terms of pooling their funds for saving and borrowing together or transactions based on their other activities such as sale of goods or services to one another. Therefore internet and mobile banking and payments system can be considered. Further, groups may wish to advertise specific products to other groups or place adverts seeking certain goods or services from other groups. E-advertising can therefore become a reality for networked groups seeking to broaden their markets. On the downside we must acknowledge that there are certain risks that could become a reality in the process of trying to actualize these opportunities and therefore need to be taken into consideration when designing solutions for linking self-help groups. Risks such as those associated with virtualization are a good example one of which is security. The security of such solutions must be enhanced so that users' information and transactions are protected from abuse and misuse. There are also social risks that could occur for example the sharing of information across cultural and ethnic divides could be misused to trigger ethnic conflict and strife.

Some broad observations on the realities of the environment in which these groups in Africa operate can also help guide interested parties in the development and design of e-solutions for group linkages:

- Cash is still the main mode of transactions and savings in most of Africa and in the case of groups this is mainly because group members prefer to see everyone's contributions and payments as hard cash which they can all verify and attest to having been paid. Mobile banking and mobile payments are nonetheless fast becoming a reality with Kenya being a good example of where the population has really embraced such services especially at individual level as they help fill the gap of limited traditional banking services. Usage of such services within groups therefore has potential to grow. Virtual payment systems are thus a possibility for use within an e-solution for internal and external group transactions. However this would need to be designed in such a way as to capture the trust of users for example confirmation messages and receipts for all transactions carried out through the solution may go a long way in improving the user experience in terms of trust of virtual systems. The system would also need to be fast enough so that once a transaction is completed online, the users would be able to contact the party they are transacting with and confirm that indeed they have also witnessed the same confirmations on their end and that all is in order. Use of PINS to secure personal financial information and transactions would also be necessary. (The African user would likely be more conversant and comfortable with the use of PIN numbers as they normally use these to secure their SIM cards on their mobile phones as opposed to passwords which are commonplace with web based applications).

- Informality and self-management is common in most types of self-help groups particularly ROSCAs while in some others such as ASCAs there is a little more structure. An e-solution designed for use by self-help groups would therefore need to support this kind of informality; groups should be able to conduct their internal activities as well as network with other groups in an ad-hoc manner on the platform so that it is as much as possible as natural as it is in face-to-face interactions. The solution should not overwhelm the users with too many formal procedures alien to their normal informal structures and way of operation for example if they have to register to use the solution, the information required of them should be the utmost minimal to enable activation.
- Groups serve multiple needs such as financial services, literacy, health, education, agriculture, business support, social welfare and support. They also tend to evolve over time in terms of functions and activities so as to cater for changing needs of the groups and their members, and the changing environment. Assuming a web platform therefore as the e-solution designed for group linkages, it would auger well to have a variety of features and functionalities to cater for the different group needs and activities and in conjunction with this the possibility for groups to select the relevant ones for their group. Tracking capabilities for instance of usage of the various functions of the e-solution would be useful as well as rating and feedback systems of the solution so as to enable the developers keep abreast of the changing needs hence evolve the solution alongside or in anticipation of the group evolution so as to remain relevant to the users.
- In rural Africa there are still literacy challenges both in terms of formal education as well as use of technology. The kind of user targeted for such e-solutions tends to fear interaction with technological or computer based systems; they think that because of their inexperience any attempted use by them will result in breakage or damage of the tool or system. It may be necessary at the onset to design the solution to be as easy to use as possible perhaps through simple icons and local imagery to aid navigation as opposed to text menus. Where menus are used they ideally should be very basic use menus that are extremely easy to learn and to use so as to ease the users into the interaction. With frequent use they are likely to get comfortable and more adept to advancing to more ‘complex’ systems. Heavy use of graphics, pictures, videos and other relevant media also needs to be enabled on

the platform for purposes of content so as to better support learning and knowledge exchange between groups. The solution should also be built to support group usage and learning as opposed to individual use.
- There are inherent infrastructural challenges in many parts of Africa such as lack of electricity, poor roads and road networks, and limited access to computers. Any e-solutions for group linkages would therefore need to be designed as a low energy use solution or one that can make use of renewable energy technologies; limited access to electricity would therefore not hinder active use of the solution.
- Most groups keep small fund sizes and this is also due to low economic capability of their members. The solution so designed needs to be as low cost as possible in order to be affordable to the low income users who are the majority.

The ideal solution for linking the self-help groups therefore would be one that educates the group members by providing access to literacy skills, general livelihood skills and by increasing social and cultural interaction and education. It should also be an enabling one that gives the users access to financial services such as savings, investments, loans and insurance, and also elevates the group members and the group as a whole in their community by allowing them to be more visible to government, NGOs, private sector and provides a platform for them to receive development interventions.

The use of mobile telephone technology is a cost effective opportunity for linking groups that may be geographically dispersed. In Kenya for example, the availability of service provider networks and handsets in almost all areas of the country make mobile telephony a potentially cheaper and convenient mode of networking. It was noted in the study that many households had access to at least one mobile phone within walking distance in their villages. This is a strong motivation to create opportunities for linkages via the mobile phone network. Enhanced services from service providers such as money transfers, credit on talk time, sharing of talk-time units and call back services have promoted mobile phone usage. With the advent of mobile telephony, radio has also become more interactive and an important source of information that has been responsible for making a difference in the livelihoods of many community members. The potential for disseminating content that meets the needs of the community either through mobile handsets with a radio feature or opportunity to *call-in* to clarify a matter or express an opinion on a radio programme are becoming popular. In addition, given the limited access to computers, it is very likely that most people in Africa will have their first experience of the internet through mobile phones and are likely to continue accessing the internet through their mobile phones. Informal self-help groups and other community based organisations are very potential first adapters of mobile internet in Africa given the opportunities it may present to them such as a platform for sharing and exchange with other groups as discussed in this paper.

6 Conclusion

Limited access to information by informal self-help groups requires innovative ways other than the conventional means which have been overstretched by increasing populations in the developing world. At the end of the day, the process of linking self-help groups would mark the beginning of an enrichment of the individuals in groups as well as the groups as functional units in the society. A social forum or platform where groups can begin to interact introduces various opportunities and benefits for the participating groups such as access to more resources, enhanced learning, socio-cultural integration and overall increased exposure. This is so long as the created platform can be such that there is acceptance and ownership by the groups and it also fosters trust between the groups bearing in mind all the challenges and realities of groups and life in Africa as discussed. Currently most African countries have poor infrastructure that would be needed for successful group linkages for example telecommunication network access is still a challenge in some interior rural areas where many groups are located. Radio and TV signals are also not available in such areas. Access to internet is also quite limited due to limited PC access as well as the related connectivity cost which despite reduction over time is still unaffordable for majority of the rural folk. In addition general literacy as well as computer and overall technology literacy is a challenge that is yet to be fully overcome. E-solutions as demonstrated by this paper stand to offer a way forward for networking groups including the incorporation of mobile phones as access tools to the network. The research already done highlights gaps and challenges which can be translated into many and varied opportunities for design of e-solutions for self-help groups in Africa.

Acknowledgments. Various organizations partnered to ensure the success of the in-depth research study upon which this paper is based. We would like to thank Nokia Research Centre, Africa for carrying out the study together with their university partner, Institute for Development Studies – University of Nairobi and local and international development organization partners namely Decentralized Financial Services (DFS), and International Media Support (IMS) who dedicated their time and intellectual resources for the research.

References

1. Nokia Research Centre, Africa. Merry Go Round: A Study of Informal Self-Help Group In Kenya. Nokia Research Centre, Africa, Nairobi (2009)
2. Financial Sector Deepening Trust Kenya (FSD-Kenya), FinAccess National Survey. Financial Sector Deepening Trust Kenya (FSD-Kenya), Nairobi (2006)
3. Gugerty, M.K.: You Can't Save Alone: Testing Theories of Rotating Savings and Credit Associations in Kenya. Title of proceedings unavailable, Washington, USA, pp. 3-4 (2003)

User Studies on Mobile Ticketing

Jussi Rämänen[1], Sirpa Riihiaho[1], Mikko Erkkilä[2], and Antti Seppälä[2]

[1] Aalto University School of Science, Department of Computer Science and Engineering, PO Box 19210, FI-00076 Aalto, Finland
{jussi.ramanen,sirpa.riihiaho}@tkk.fi
[2] Tieto Corporation, Aku Korhosen tie 2-6, PO Box 38, FI-00441 Helsinki, Finland
{mikko.erkkila,antti.seppala}@tieto.com

Abstract. This paper presents the results of user studies on a public transportation mobile ticketing concept. The studies aimed at obtaining a thorough understanding of the diverse user needs in a travel ticket ecosystem, thereby providing relevant and valuable information to guide the development process of a new mobile ticketing system. The user studies were conducted in three phases, starting from analysis of the present ticketing system and ending with evaluation of the new concept generated. The studies were executed in close collaboration between academia and industry, and the results proved to have great impact, as they were used for making strategic choices between technological alternatives and in marketing the development of new mobile financial services to various potential business partners.

Keywords: User-centered design, User research, Scenarios, Mobile services, Public transportation ticketing.

1 Introduction

Designing a new ticketing system for public transportation is a notable challenge, as multiple user groups and their varying needs have to be addressed. Not only do the passengers use the system on a daily basis; rather the drivers and ticket inspectors do as well. Although all user groups share the same general demand for ease-of-use, each user group also has specific needs for the ticketing system. In addition to the needs related to the daily use of the system, the ecosystem also involves other needs related to the management, maintenance, and business aspects of the system. In current rapid product development cycles, it is very difficult to address and fulfill all the various needs and satisfy all the parties involved. Therefore, selecting effective methods and a well-balanced sample of representative users for the user studies is especially important. Effective communication between involved stakeholders throughout the development process is also essential to enable changes in the development early enough to save both money and time.

In recent years, there has been a noticeable trend to bring out more and more services for mobile use. One potential reason for this is the rapidly increasing number of mobile cellular subscriptions, which was globally expected to reach five billion in 2010 [1]. Public transportation ticketing is one application area that can surely benefit

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 630–639, 2011.
© Springer-Verlag Berlin Heidelberg 2011

from the possibilities provided by mobile technologies. Besides, as public transportation concerns a mass of people, a successful introduction of mobile ticketing can further create opportunities for other mobile services.

This paper describes a case study where user studies were conducted to obtain a thorough understanding of the diverse user needs in a mobile travel ticket ecosystem. The objective of the study was to form a strong basis for the development of a Near Field Communication (NFC [2]) -based mobile ticketing concept. Before this case study, the development project had concentrated primarily on technological aspects, so the need for a user-centered approach was evident. Therefore, to broaden the scope beyond technological feasibility prototyping, user studies were conducted to get ideas concerning how to make the system easy and pleasant to use and even tempting for the passengers. Although the focus was on the passengers on public transportation, drivers and ticket inspectors were also included in the studies. Furthermore, the public transportation company and the developers of the planned system were involved in the studies through several workshop sessions.

In section 2 of this paper, we present the background information of the project and the technological framework, as well as the parties involved. The third section describes the process of the study; following this, section 4 gives an overview of the applied methods. Section 5 elaborates on the results of the study, whereas section 6 discusses the experiences we had in the case study and the role of our user studies in the continuum of system development. Finally, section seven presents our conclusions.

2 Background

The case study in this paper is a part of a Mobile Financial Services (MoFS) research program funded by TEKES (the Finnish Funding Agency for Technology and Innovation). The aim of the program is to promote the transition of financial services to the mobile environment, and to steer present use habits toward a “wallet in a mobile phone” mindset [3]. Both academia and industry are involved in the program. In this case study, Aalto University School of Science and the Tieto Corporation (a Finnish company offering IT, R&D, and consultancy services) were the most important parties.

At the moment, the public transportation of the Helsinki region is equipped with a system called Travel Card (see Fig. 1) that is based on radio-frequency identification (RFID) technology. The system consists of card readers and ISO 14443A–compliant DESFire smartcards [5]. Users can choose between two ticket alternatives: value and period. Value ticket users purchase single tickets by showing their card to the reader and pressing one of its numbered fare zone buttons, while period ticket users only show their card to the reader [6]. A Short Message System (SMS)-based scheme is also available, but with limited coverage of the transportation ecosystem.

The starting point for the development of the mobile ticketing system was to utilize NFC technology: Instead of a separate card, the tickets would be handled by an NFC microchip inside a mobile phone. Like RFID, NFC is a technology for short-range wireless connectivity that can also be used for peer-to-peer connections and reading passive RFID tags [2].

Fig. 1. Travel Card [4]

As we planned the user studies, we identified at least three challenges that would have to be addressed. First, we should be able to present feasible and inspiring ideas that would meet the requirements of both the development team and potential business partners, while maintaining a strong focus on the user perspective.

Secondly, the public transportation ticketing system has multiple user groups. In addition to passengers, bus drivers and ticket inspectors have their own perspectives and needs regarding the use of the system. These user groups are also known to be harder to contact than ordinary passengers are.

Thirdly, obtaining relevant user comments on concepts was not obvious, because of the unfamiliar NFC technology framework. This had to be taken into account when presenting ideas to the users.

3 Process

The user studies were conducted in three phases, namely:

1. Exploration of the current ticketing system and previous studies;
2. Idea generation to create scenarios of mobile ticketing use cases and value-adding services;
3. Evaluation of the scenarios with passengers.

The first phase of the user study aimed to construct a comprehensive understanding of the present ticketing system. The material from this phase served as a basis for a realistic context of use and the key requirements for the new system. The methods used in this phase included a survey, observations, and interviews with various user groups (passengers, drivers, and ticket inspectors). We also used previous studies (e.g., [7]) and student assignments as supporting material. A heuristic evaluation [8] was conducted at the end of this phase to support the classification and rating of the usability problems found.

Based on the results of the first phase, we started to generate ideas on the functionalities and value-adding services of the mobile ticketing system. Ideas were kept independent from the technological solutions, and the focus was on the passengers' interaction with the system. In the ideation sessions, brainstorming [9] and 6-3-5 brainwriting [10] were utilized. To illustrate some potential use cases of the mobile ticketing system, we drafted a few scenarios that were presented in a workshop session with representatives of both the potential customer (i.e., Helsinki Region Transport) and the developers of the system (i.e., Tieto).

Based on the workshop feedback, we formed one comprehensive scenario that covered the whole ticketing lifecycle from the passenger's point of view. The scenario started from making the user aware of the service and motivated to download the application to her mobile phone, and ended with the actual use of the application. In addition, some value-adding services were presented in the scenario to encourage the users to share their ideas about useful supplementary services. The finalized scenario was evaluated with 21 passengers of varying backgrounds. At this point, the drivers were not interviewed due to their tight schedules. Finally, the collected feedback was analyzed and presented to all the major public transportation organizations in Finland. It was also used as a basis for discussion in several workshop sessions, and as a marketing material in events promoting mobile financial services to other business partners such as telecom operators and banks.

4 Methods

In our user studies, we utilized various methods, namely surveys, observations, interviews, heuristic evaluation, and scenarios to collect both quantitative and qualitative data. Quantitative data was considered important in convincing the stakeholders of the severity of the usability problems in the current system and, thereby, of the need for a new system. Qualitative data, on the other hand, was considered essential for obtaining in-depth information on how passengers would like to use a mobile ticketing system, problems and threats they might attach to it, and supplementary services they would like to have available. In the next four subsections, we will elaborate on the methods that gave us the most valuable results.

4.1 Surveys

Surveys can be used to collect considerable amounts of user data quickly and with little effort. In this case study, we conducted a web survey, for which we gathered participants mainly via the Helsinki Region Transport website. In two weeks, we received 162 responses. The survey was conducted at the very beginning of the study to get a wide perspective on the passengers' opinions and needs as early as possible. In the survey, we concentrated on gathering data on user experiences of the present ticketing system and ideas on how a mobile phone could be used for public transportation ticketing.

4.2 Observations

In observations, the observer follows the actions of one or more people while they are carrying out their tasks. People can be observed from a distance without interaction, or by getting involved and interacting with them. The non-intrusive observation technique is especially suitable in cases where the involvement can cause unwanted changes in the dynamics of the situation, whereas involvement can sometimes reveal important subtleties.

We carried out observations in two stages. The aim of the first round of observation was to obtain qualitative data, while the second observation round concentrated on recording quantitative data. We collected qualitative data by observing the use of the current ticketing system from different perspectives, focusing on the factors that seemed to affect the use most significantly. The collected material was used as a basis for forming a comprehensive context definition covering differences in the means of transportation, the different user groups and their roles, and the effects of environmental factors. The context definition served as an important source of information for the requirements of a new mobile ticketing system.

In collecting quantitative observation data, the focus was on identifying usability problems and collecting data on their frequencies. During the second round of observation, we collected over 5 hours of observational data on 315 passengers entering busses, trams, local trains, and the subway. The quantitative observation data served as means to justify the need for a new system to the representatives of the public transportation companies. All of our observations were non-intrusive, since the aim was to get a realistic view on the problems that passengers face while entering the vehicles.

4.3 Interviews

Interviews can be used to obtain in-depth information about specific themes. Depending on the objective, the interview can be based on closed questions, wider themes, or a mixture of both. The interviews we conducted during the case study were semi-structured, avoiding closed questions and focusing on interviewees' comments and opinions on a few specified themes.

At the beginning of the study, we conducted four individual interviews and one group interview. Two bus drivers, one tram driver, and one ticket inspector participated in the individual interviews, and five passengers participated in the group interview. We interviewed the drivers and the ticket inspector in their real work environment. The real context, with all the tools and systems at hand, was considered to give valuable results, since it would remind the drivers and the inspector of situations that were relevant and descriptive for the Travel Card system. The aim of the passenger group interview was to expand on some of the findings from the survey responses, such as flexibility in time and place in loading the money or period onto the card.

In the evaluation phase, we interviewed 21 users individually about our finalized mobile ticketing scenario. Our aim was to gather information about motivational factors, perceived threats, and potential usability problems. The interviewees included 12 women and 9 men whose ages varied from 15 to over 50 years; most of them had a travel card loaded with either a period or value.

4.4 Scenarios

Using scenarios in the design and evaluation of new systems is an effective way to convey concept ideas in a format that is evenly accessible for all of the stakeholders, regardless of their background [11]. We used scenarios to illustrate our ideas for various stakeholders involved in the project. Our first scenarios of how passengers would use the system were narrative descriptions supplemented with sketched comic strips. The finalized scenario was constructed around a series of computer-edited photos from different phases of the imagined use of the concept system (see Fig. 2). Familiar devices such as a regular mobile phones and the current card reader placed in a realistic environment were considered to reduce confusion caused by unknown technology and encourage users to express their ideas and opinions.

Fig. 2. Examples of scenario visualizations

5 Results

As we used various methods in our studies, we managed to cover a wide variety of situations and different factors. In this section, we present the most important results of the case study. The results have been sorted in terms of the data-gathering method.

5.1 Exploration of the Current Ticketing System

Based on the survey results and the group interview session, it was evident that possibilities for checking the travel card value and validity were currently too limited. Checking the information is possible only with a card reader device and in some of the ticket vending machines. The passengers wanted to check their card value and validity more flexibly, irrespective of place and time. They also wanted a more flexible alternative for loading value and period onto the card, since this is currently possible only in kiosks and in some of the ticket vending machines, which might not be at hand when needed. During our user studies, we also discovered that the passengers thought the card reader did not provide enough feedback on the actions, showed the feedback inconveniently, and did not allow for correcting false button presses that the passengers occasionally made while trying to ensure that they had purchased the ticket successfully.

The recorded observation data gave good insight into the problems people face while using the system. We discovered that every other user had some difficulties in purchasing a value ticket with their Travel Card. It was also noticed that people mostly used both hands to select the fare zone on the card reader, so they had to lay down their bags, thus slowing down the boarding of the bus. In problem situations, the bus drivers were not always willing or able to help the passengers. The observations gave valuable information on the requirements of a mobile system for reducing problem situations. For instance, the mobile ticketing system should be easy to use with one hand, as well as if the user is on the phone while entering the vehicle. The quantitative results of the observations gave credibility to the requirements that we presented in our workshop sessions.

In the individual interviews with the drivers and the ticket inspector, the need for fluent and reliable use of the ticket system was emphasized. For example, the bus drivers mentioned that they often needed to help people to select the right fare zone, which slows down boarding. The drivers also pointed out that the cold winter climate causes problem situations in reading the smartcard. The interviews also revealed that the drivers and the ticket inspectors did not like the current SMS system at all, since it was difficult to validate the tickets from a tiny phone screen, especially in poor lighting conditions. The interviewees emphasized the need for a system that enables easy validation of tickets and, therefore, thought that NFC-based mobile ticketing could suit their needs well, since it resembled the current travel card system in terms of the ticket validation procedure.

5.2 Evaluation of the Scenarios with Passengers

As we evaluated the mobile ticketing concept with passengers, we received very positive feedback about the fluency and flexibility provided by the mobile system. Flexibility was considered the most important motivational factor for adopting the mobile ticketing by almost half of the interviewees. Another significant benefit of mobile ticketing, highlighted by approximately a fourth of the users, was the fact that their mobile phone was something they almost always had with them, whereas a separate card was forgotten at home more often as a consequence of using different jackets or handbags. Users also noted other positive aspects of the mobile system, such as the possibility of useful supplementary services. However, value-adding services were not by themselves considered reasons for switching to mobile ticketing. Our interviews also revealed that while receiving monetary benefits was generally a tempting idea for most people, it was not considered a major motivational factor for a change of service; rather, it was more something that would increase satisfaction.

Although the users considered the idea of mobile ticketing to be mostly beneficial for many reasons, they also expressed some concerns. Many of the users were worried about the security of the system, as well as the potential problems in downloading and setting up the service application. In the scenario interviews, we provided the users with two alternatives for accessing the service, and asked which one they preferred. The alternatives were 1) downloading the mobile ticketing application by placing the mobile phone near a passive NFC tag in a bus stop advertisement, or 2) sending an SMS to a provided number and receiving a hyperlink to download the application. Many of the interviewees perceived the first alternative as a relatively unsecure way to download the application because of the potential risk for viruses and malware. However, it was generally considered a convenient way to start using the service. The latter alternative, utilizing SMSs, turned out to be more widely accepted, as users thought they would have more control over the process, thus reducing the security risks. Regarding both of the alternatives, the users were concerned with data transfer costs, as the majority of the users did not have a broadband Internet connection with fixed pricing. Many of the users also pointed out the problems that they had experienced when trying to install applications to their mobile phones.

In addition to the previous concerns, users also expressed their suspicion related to some other mobile-specific matters. Lack of technical reliability was seen as one threat, since new advanced smart phones were considered relatively unreliable. Furthermore, some of the users mentioned running out of battery as a possible problem. The current smart card solution was considered more reliable in these terms.

In addition to the weaknesses of the existing ticketing system that were uncovered, the results of our study highlighted important facts about users' attitudes toward and expectations of the mobile ticketing system. One remarkable finding from the user interviews was that a significant proportion of users expressed their unwillingness to buy a new NFC-enabled mobile phone just to be able to use a mobile ticketing system. Some stated that they might consider NFC compatibility when purchasing a new phone, but it was not seen as the determining factor in the purchase decision.

6 Discussion

In our user studies, we noticed that the results from different methods clearly highlighted different aspects that effectively complemented each other. For example, the interviews with the drivers and the ticket inspector revealed weather-related problems that were not noticed during our observations in the summertime. In addition, involving multiple user groups brought out varying needs. While the drivers were interested in the efficiency and fluency of boarding the vehicle, the ticket inspector focused on ergonomic aspects and simplicity in relation to checking the validity of tickets.

The quantitative results proved to be valuable in convincing the representatives of Helsinki Region Transport of the severity and frequency of usability problems with the current ticketing system. Although the problems with the system were known, the representatives were stunned by their frequency.

Scenarios turned out to be a natural way for all the stakeholders to understand and comment on the context and usage of the projected system. Using real photos enabled us to build a comprehensive and yet credible and realistic scenario quickly and without the need for good drawing skills. As the use context and technological devices, such as the mobile phone and card reader, were familiar to the users, they could easily concentrate on the new service, and assess its utility and usability.

The whole process of user studies lasted less than two months. Nevertheless, we were able to gather diverse material, helping us to form a holistic view on the various issues involved in the ticketing ecosystem. We were also able to address the major requirements of the various stakeholders, and came up with solutions and ideas that were well accepted by both the users and the company representatives.

The studies revealed that the users had interesting prejudices, as well as attitudes and habits that were, to some extent, in conflict with the developers' assumptions. For example, the average phone renewal cycle turned out to be considerably longer than assumed. Due to the lack of NFC-capable devices on the market and the users' unwillingness to buy them, strategic decisions were made in the development process, and new solutions had to be considered. From this perspective, it was determined that the shift to an entirely mobile-based ticketing solution would be more successful if handled through smaller steps.

From our user studies, flexibility turned out to be the most important factor in user expectations of the mobile ticketing system, as the current system was considered too rigid. Our studies also revealed that a majority of the users were unwilling to enter extensive amount of information with a mobile device due to a limited display area and a small keyboard. Two-thirds of users wanted to limit the use of mobile phones to setting up the service and daily use, and preferred to give most of the registration information with a desktop computer. Based on these findings, a decision was made to provide a partially mobile ticketing solution as the first step toward a completely mobile system. In this first step, some features such as loading period and value would become available through mobile phones and desktops, but the daily use in the vehicles would still be handled in the current way. This approach would bring flexibility to the system and familiarize the users with mobile services in the public transportation context while preserving the familiar everyday use logic.

7 Conclusions

Although interest in mobile services has increased substantially, and users in our studies saw the advantages of the mobile system over the current system, there are still issues that have to be considered. Based on our findings, security concerns and other attitudinal matters seem to be significant factors affecting the user acceptance of mobile services in the context of public transportation. Additionally, the users seemed unwilling to put much effort into the adoption of a new ticketing system.

The results presented in this paper have given an overview of the problems with the current ticketing system in the Helsinki region, as well as users' attitudes and opinions concerning a mobile ticketing solution. It may be that not all of our results are directly applicable to other public transportation ticketing ecosystems, but the results give insight into the attitudes various user groups have related to this kind of service. Most importantly, we believe that the results presented highlight the different contributions of the methods utilized, thus giving ideas for methodological choices in similar cases. Our case study also serves as an encouraging example of how versatile and elaborate user studies, combined with effective communication throughout the process between the stakeholders involved, can support decision-making in the development of a new system.

Acknowledgments. We want to thank Riikka Hänninen, Satu Raudasoja, Hanna Salmén, and Tuomas Vaittinen for their excellent contribution in the user studies, Professor Marko Nieminen for his supervision and effort in creating the survey, and Mari Tyllinen for suggestions for improvements.

References

1. International Telecommunication Union: ITU Sees 5 Billion Mobile Subscriptions Globally in 2010, http://www.itu.int/net/pressoffice/press_releases/2010/06.aspx (retrieved January 20, 2011)
2. NFC Forum, http://www.nfc-forum.org/home/
3. MoFS - Mobile Financial Services, http://mofs.soberit.hut.fi/
4. Helsinki Region Transport: New Travel Card: What Is It All About?, http://www.matkakortti.net/en/new-travel-card/what-is-it-all-about/
5. Helsinki Region Transport: Millaiset ovat matkakortin tekniset ominaisuudet? http://www.hsl.fi/FI/matkustajanopas/faq/Matkakortti/Sivut/Millaisetovatmatkakortinteknisetominaisuudet.aspx
6. Helsinki Region Transport: Tickets on Travel Card, http://www.hsl.fi/EN/ticketsandfares/ticketsontravelcard/Pages/default.aspx
7. Mallat, N., Rossi, M., Tuunainen, V.K., Öörni, A.: The Impact of Use Context on Mobile Services Acceptance: The Case of Mobile Ticketing. Information & Management 46(3), 190–195 (2009), doi:http://dx.doi.org/10.1016/j.im.2008.11.008
8. Nielsen, J.: Enhancing the Explanatory Power of Usability Heuristics. In: Proc. SIGCHI Conference on Human Factors in Computing Systems CHI 1994, pp. 152–158. ACM Press, New York (1994), doi:http://doi.acm.org/10.1145/259963.260333
9. Rossiter, J.R., Lilien, G.L.: New “Brainstorming” Principles. Australian Journal of Management 19(1), 61–72 (1994)
10. Lumsdaine, E., Lumsdaine, M.: Creative Problem Solving: Thinking Skills for a Changing World. McGraw-Hill Inc., New York (1995)
11. Carroll, J.M.: Introduction: The Scenario Perspective on System Development. In: Carroll, J.M. (ed.) Scenario-Based Design: Envisioning Work and Technology in System Development, pp. 1–17. John Wiley & Sons, Inc., Toronto (1995)

Photo Diaries – A Peek into a Mobile Worker's Life

Anthony C. Sampanes, Michele Snyder,
Lynn Rampoldi-Hnilo, and Brent-Kaan White

Cisco Systems, Inc., 170 Tasman Dr
San Jose, CA 95134
csampane@cisco.com, michele.snyder@oracle.com,
lynn.rampoldi-hnilo@oracle.com, brent.white@oracle.com

Abstract. Understanding mobile users and their behaviors in context, across multiple countries and cultures, is challenging and costly. However, the reward of creating a mobile application that meets mobile users' needs (and is actually used - not deleted within days of being downloaded) is priceless to the user in terms of productivity, the employer by keeping data updated, and the developer by producing a successful application.

To develop such applications, mobile HCI professionals need to identify mobile users' daily mobile habits and tasks completed with their devices. To uncover unexpected uses and a range of contexts, we recommend using a photo diary technique in conjunction with other ethnographic methods. Conducting mobile photo diaries will help mobile HCI professionals in their application decision making process, by giving them additional insights into their users that may have been inaccessible or unthought-of before, and visually rich deliverables to share with management.

Keywords: Diary, Photographs, Ethnography, Research, User-Centered Design, Mobile.

1 Introduction

HCI researchers need to look at cultural differences in behavior because many products today are developed for a global market. National laws, cultural backgrounds, socio-economic backgrounds, religion, and language can all affect how people interact with a product in different countries [1]. The rise of global products has created new groups of users that we need to understand more clearly. Mobile users fall clearly into this category. Mobile application adoption is extremely sensitive to cultural demands and requires an understanding of each culture one is going to ship their product to. Getting an insider's view, a documentation of what folks are doing with their mobile devices in context, is crucial to understanding their culture and adoption of it. The photo diary method lends itself nicely to collecting this valuable data. More so, since studying mobile behaviors often includes travel and a number of locations, this technique allows mobile HCI researchers to gather information and insights that are generally hidden from us. While diary studies have been frequently used in HCI [2, 3, 4, 5, 6], they rarely have been used to assess cultural differences among users.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 640–647, 2011.
© Springer-Verlag Berlin Heidelberg 2011

1.1 The Diary Method

Diary studies are an established method [7] that can be used for capturing all kinds of information, from shopping behavior [8] to the behavior of business professionals [9] to identifying interesting uses and contexts of mobile devices. Most commonly, this method has required participants to capture (through written descriptions, audio recordings, and pictures) information about experiences that happen during the day. Diary studies might require participants to capture data at particular times, specified time intervals, or when particular events occur [10, 11, 12]. Diary studies can range from relatively unstructured accounts, with some blank lines to write in information [13], to highly structured accounts, with detailed activity checklists [14].

Diary Method Advantages. Diary studies are advantageous to other methods of data collection for many reasons. Diary studies allow information to be captured from the participant's perspective, in their own form of communication, and in the context of the real-world. Information that can be hard to gather because it occurs infrequently, is socially or contextually restricted, or is difficult or impossible to observe can easily be collected if the diary is oriented appropriately. Furthermore, because the participants are generally explaining their behaviors or experiences in the diary, the diary enables researchers to gain a deeper perspective on the impact that situations or events have on participants.

Diary studies can be extremely useful in examining mobile and cross-cultural issues for a multitude of reasons. Diary studies are inexpensive to run compared to other methods. Costs are minimal compared to the high costs of international travel and accommodations (often necessary to collect data using other methods). Because other methods are generally more expensive, their associated data collection phase is often shorter and involves soliciting data from fewer locations (that is, countries). For instance, a field study may occur only for a few hours per day, for a few days, and in only one country. A diary study, however, can more easily be conducted over a longer period of time, in more countries, allowing for a wider and more diverse set of data to be collected.

Diaries also enable researchers to gather data at time intervals that are often potentially difficult to capture using other methods. For example, when a person wakes up she may perform a task, such as making a call during breakfast, which could be of interest. Another example may be a task like watching a video late in the evening before going to bed. These are potentially times when important differences in behavior may occur. This is especially true in the mobile space, because mobile devices are used ubiquitously (as they can easily be transported from place to place and in many cases remain at person's side 24 hours a day). By collecting data at these different time intervals and environments, researchers can capture data from the cultural perspective of the participant, without the interpretation of a researcher who may not be from that culture.

Lastly, a diary study tends to be less intrusive because participants can typically fill out a diary on their own time at their own comfort level. This can make it easier to recruit a larger number of participants.

The Photo Aspect. Diaries that include taken photos are especially good for HCI professionals studying mobile users from different cultures. Most present-day phones have built in digital cameras. This allows for easy data capture while a person is using their mobile phone with little interference of their mobile use. Furthermore, photos capture a lot of information and often more clearly and fully document places and cultures. This comprehensive account becomes especially important if the cultures or locations are places where a researcher can't follow or is relatively unfamiliar (such as a different culture). Visual deliverables, such as pictures, serve as illustrative and compelling data (compared to written documentation [15]). They can also be extremely informative when explaining one culture to another. Miscommunications between different cultures happen frequently. When photos or other visual aids are shared in conjunction with verbal or written descriptions, miscommunications are less likely to occur.

1.2 International Mobile Study Objectives

The present photo diary study was one part of a larger mobile, international ethnographic research project aiming to observe where, when, and how enterprise workers use their mobile devices and to understand key aspects of the mobile culture. We particularly wanted to get participants' perspectives on certain aspects of mobile usage (rather than the view of a researcher looking in on participants). We also knew that we would be limited in the times and places that we could follow participants; wanting to collect a robust set of data, we needed to expand the scope of places and behaviors captured. This was particularly important because a mobile device is with users nearly all the time, in nearly every location. Finally, we wanted to collect pictorial data to contribute to visual deliverables that we were planning (such as personas, culture boards, and a calendar). This case study will be presented to show how a photo diary was used to gather cross-cultural data.

2 Procedure

The photo diary was administered in conjunction with a larger international ethnographic research project, conducted in India, Singapore, and the U.S. The entire study aimed to (1) understand users across technologically advanced cultures, diverse mobile workforces, and emerging markets and to (2) observe where, when, and how mobile workers used their mobile devices. The photo diary was administered to support those goals and to accomplish the following:

- Collect pictorial data to help create visually based deliverables
- Get participants' perspective and personal insights
- Expand the scope of places and behaviors observed

Twenty-four of 33 people given the diary returned it via e-mail, postal mail, or during a subsequent meeting with the researchers. The participants represented seven different business roles from India (n = 6), Singapore (n = 11), and the U.S. (n =7).

Participants were either e-mailed a soft copy of the diary or were handed a paper copy during an associated follow-along activity. In paper form, the diary was 10.8

centimeters wide, 14 centimeters tall, and less than ½ centimeter thick, so the diary was small enough and light enough for participants to easily carry around with them wherever they went. Participants were asked to complete the photo diary over a week's time. Completing the diary required participants to document mobile activities about scenarios predetermined by the researchers. Participants documented such activities by taking pictures (either with disposable digital cameras given to the participants by the researchers or with their own personal cameras) and providing written descriptions to support the pictures. The written descriptions documented information about the activities themselves, the environments, and any other necessary information about the context of the tasks.

The diary scenarios focused on where the participants used their mobile devices (for example, the most crowded places and loudest places), how they used their mobile devices (for example, the tasks that they performed), challenges they faced while using their mobile devices (for example, the most difficult place to use their mobile devices), and work versus personal usage (for example, the most common places used to complete work tasks and the most common places used to complete personal tasks).

Data came in two forms: 1) photos taken by the respondents, and 2) the associated short answer responses that described the picture, context, and mobile activities (see Figure 1). The data was examined by tallying frequencies of common responses and by identifying common themes. Themes were focused around common responses, common places, common tasks, and common challenges.

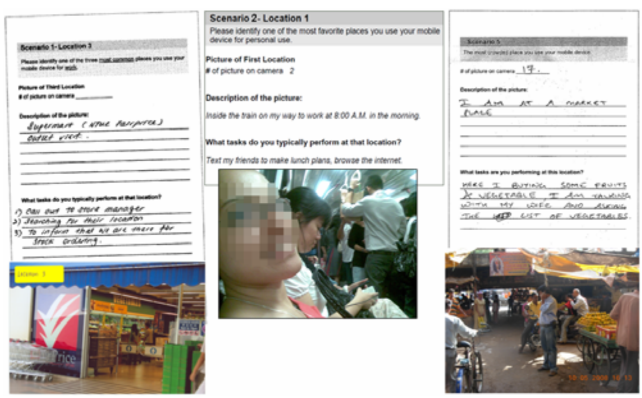


Fig. 1. Examples of diary entries

3 Findings

This study yielded rich results that went beyond the original goals of the study. The photo diary method provided us with four types of findings that would have been difficult or costly to obtain with other methods.

3.1 Information Normally Observed

The first type of finding was a private and personal perspective that captured information about places and times where we could not typically observe a mobile participant. For instance, we found that bathrooms were a place where people heavily used mobile devices across all three countries studied. Many participants provided information about how they used mobile devices while in a bathroom; a lot of pictures of bathrooms or toilets were submitted. Another example is the wealth of information about their bedrooms and living quarters. Of note was the frequency of use of mobile devices within the home. We had not considered this; we thought that mobile was when people were on the move, etc. Pictures were extremely useful for this project because mobile behaviors are often subtle, quick, or small, which makes them challenging to capture when a researcher is observing a participant in the field.

3.2 Cross Cultural Similarities

One of the goals in creating the personas was to highlight cross-cultural differences. The second type of finding gained from this study was the common mobile themes across the cultures. We found similarities across the cultures regarding the type of places that people use their mobile devices, along with some reasons why mobile devices were used. Much of the work life among the cultures was similar, including work tasks. Pictures were extremely useful because they could be spread out, making it easier for researchers to scan and detect common themes in the data. These themes were then validated by further analysis.

3.3 Cross Cultural Differences

The third type of finding illustrated the differences among the cultures. For example, some methods that were used to accomplish particular tasks, such as gathering information about business locations, varied. India varied from the U.S. and Singapore in that people would SMS to particular services the name of the business they were looking for and the service would send a return SMS with the address. In the U.S. and Singapore participants were much more likely to browse the web for that information. Pictures were valuable in helping detect subtle differences among the cultures. For instance, respondents indicated that they used their mobile device when buying things (for example, to call and ask his wife a question). The pictures showed differences in the types of locations where people purchased items in the different countries (for example, in outdoor markets, stalls, and stores). The pictures also revealed valuable data about the infrastructures and environments of the three countries. For example, Indian participants often provided pictures that portrayed chaotic streets with cows and goats intermingled with a cars and a high density of people. The pictures taken in the U.S. and Singapore revealed more organization and a lower density of people.

3.4 Validation of Ethnographic Research Findings

The fourth type of finding was supplementation and validation of the following-along data we collected as a part of the larger ethnographic research project. The photo diary supplemented our other observational data by providing times and places where we didn't observe and by providing a participant's eye-view on the data. The photos and details captured in the photo diary also served to validate a lot of conclusions we reached from observing the participants.

4 Application of Findings

One of the most important issues in the collection of the data is the ability to present and convey the findings in a way that can be used and disseminated through an organization. While Word, Excel, and PowerPoint can create a means of documenting a research project, they are basic deliverables at best. We wanted to take the visual nature of the photo diary data and showcase the wide range of different pictures and information in a variety of visual deliverables targeted at various audiences, ranging from product development to executives, to maximize the impact of the data. The pictures taken by participants worked well for these deliverables because they had a sense of authenticity about them that we did not get from photos taken by us during the follow along research. Pictures, themes, and quotes were heavily used in developing three types of deliverables: personas, culture boards (visual representation of mobile culture), and a mobile research themed calendar.

The pictures were used in the creation of a set of mobile personas, which developers and product teams use to help them build mobile applications for a global audience. The persona provided a way to capture information on mobile work environment, mobile tasks, and a day in the life using a mobile, demographics. One of the primary goals of the personas was to help product teams understand the cultural similarities and differences in how people accomplish mobile tasks across different countries. This in turn will help ensure that teams develop successful products that can be used across many cultures. The pictures from the diary study were excellent in conveying a lot of cultural information and added a very personal perspective to the personas.

Many of the diary pictures were also used in the creation of culture boards. Culture boards are large-scale displays used to elicit the feeling of a particular place and people, typically through pictures, facts, quotes, and keywords. A separate culture board was created for each county visited, and each culture board contained information on the mobile infrastructure, phone-buying habits, food, entertainment, landscape, and work environment. The diary photos were particularly good at capturing and displaying information about tasks that we had not gathered with other ethnographic methods.

Researchers created a 13-month calendar using many of the pictures collected. The calendar focused on key mobile findings and themes found in the research. The visual appeal of the calendar relies partially on its daily repetition, with periodic changes that introduce new mobile themes. Each month, a research topic is revealed within a design of a bright background, interesting pictures, study results, and succinct quotes that reinforce the research theme. Photographs from the photo diaries made the calendar visually rich and unique; they also served to initiate and maintain interest in the calendar.

Fig. 2. Snapshots of visual deliverables created

5 Conclusion

With the shift towards globalization of products and mobile devices allowing usage in previously inaccessible places, the need to understand new use cases and the similarities and differences among different cultures is increasing. Gathering this type of data can be difficult and costly. We provided a case study to show how a photo diary could be used to supplement ethnographic research. The photo diary allowed for data to be inexpensively collected in times and places where researchers couldn't follow. The diaries put a participant's eye-view on the data, allowing insight into the respondents' personal lives. Data collected through the diaries supplemented and validated the findings from the ethnography and contributed to visually rich deliverables.

While we use the photo diary as a supplemental technique, there are situations where this technique could be used as the sole means of data collection. If researchers are particularly interested in personal events (especially those that occur in times or places that make observation difficult) or want to collect end-user perspectives (as in sales or marketing), a photo diary could potentially provide all the necessary information. The photo diary can provide a snapshot that unveils findings perhaps inaccessible or unthought-of before, taking us beyond our original questions and subsequent conclusions.

References

[1] Quesenbery, W.: Making personas part of your team. Workshop presented at User Friendly 2008, Shenzhen, China (October 2008), WQusability, `http://www.wqusability.com/handouts/personas-team-uf2008.pdf` (retrieved August 20, 2009)

[2] Chin, J.P., Herring, R.D., Familant, M.E.: A usability and diary study assessing the effectiveness of call acceptance lists. In: Proceedings of the Human Factors Society 36th Annual Meeting, pp. 216–220 (1992)

[3] Palen, L., Salzman, M.: Voice-mail diary studies for naturalistic data capture under mobile conditions. In: Proceedings of CSCW 2002, pp. 87–95 (2002)

[4] Carter, S., Mankoff, J.: When participants do the capturing: the role of media in diary studies. In: Proceedings CHI 2005, pp. 899–908. ACM Press, New York (2005)

[5] Brown, B.A.T., Sellen, A.J., O'Hara, K.P.: A diary study of information capture in working life. In: Proceedings CHI 2000, pp. 438–445. ACM, New York (2000)

[6] Van House, N.A.: Interview viz: Visualization-assisted photo elicitation. In: Proceedings of CHI 2006, pp. 1463–1468. ACM Press, New York (2006)
[7] Norman, D.A.: Categorization of action slips. Psychology Review 88(1), 1–15 (1981)
[8] O'Hara, K., Perry, M.: Shopping anytime anywhere. In: Extended Abstracts Proceedings of the ACM CHI 2001 Conference on Human Factors in Computing Systems, pp. 345–346 (2001)
[9] O'Hara, K., Perry, M., Sellen, A., Brown, B.: Exploring the relationship between mobile phone and document activity during business travel. In: Brown, B., Green, N., Harper, R. (eds.) Wireless World: Social and Interactional Aspects of the Mobile Age, pp. 180–194. Springer, Heidelberg (2001)
[10] Adler, A., Gujar, A., Harrison, B., O'Hara, K., Sellen, A.: A diary study of work-related reading: design implications for digital reading devices. In: Proceedings of ACM CHI 1998 Conference on Human Factors in Computing Systems, pp. 241–248 (1998)
[11] Tavel, P.: Modeling and Simulation Design. AK Peters Ltd. In: Chin, J.P., Herring, R.D., Familant, M.E. (eds.) Proceedings of the Human Factors Society 36th Annual Meeting, A usability and diary study assessing the effectiveness of call acceptance lists, pp. 216–220 (2007)
[12] Rieman, J.: The diary study: a workplace-oriented research tool to guide laboratory efforts collecting user-information for system design. In: Proceedings of the ACM INTERCHI 1993 Conference on Human Factors in Computing Systems, pp. 321–326 (1993)
[13] Adler, A., Gujar, A., Harrison, B., O'Hara, K., Sellen, A.: A diary study of work-related reading: design implications for digital reading devices. In: Proceedings of ACM CHI 1998, Conference on Human Factors in Computing Systems pp. 241–248 (1998)
[14] Sohn, T., Li, K.A., Griswold, W.G., Hollan, J.D.: A diary study of mobile information needs. In: Proceedings of CHI 2008, pp. 433–442. ACM, New York (2008)
[15] Brown, D.: The visual vocabulary three years later: An Interview with Jesse James Garrett. Boxes and Arrows (2003)

An Experience Sampling System for Context-Aware Mobile Application Development

Jungsuk Seo, Seunghwan Lee, and Geehyuk Lee

Korea Advanced Institute of Science and Technology
291 Daehak-ro, Yuseong, Daejeon 305-701, South Korea
chaoticblue1@icu.ac.kr, rigelsh@gmail.com, geehyuk@gmail.com
http://hcil.kaist.ac.kr

Abstract. For contextual evaluation of context-aware mobile phone applications, it is desired to observe user interaction with them and collect user feedback about them in situ. For this purpose, we designed and built an experience sampling system (ESS) for mobile phones. The resulting ESS is characterized by the following features: event-driven sampling of user context and feedback, real-time monitoring of user experience, and offline tools for data analysis. We conducted user experiments to confirm that the ESS was effective for evaluating mobile phone applications. We report in this paper the design considerations and the implementation detail of the ESS as well as the results of the user experiments with the ESS.

Keywords: mobile applications, context-aware applications, experience sampling system.

1 Introduction

One of the most desired attributes of mobile phone applications is context-awareness. A mobile application without context-awareness for instance may interrupt the user when she is in an important meeting, and may make a noise when she is in a movie theater. The design of context-aware mobile application is an important issue for mobile phone developers. One of the main difficulties in the design of context-aware mobile applications is their evaluation [4, 7]. Our past experience tells that the evaluation of a design was more difficult than the design itself. In this paper, we describe an experience sampling system (ESS) that we developed for context-aware application development. The primary function of the system is to monitor the usage of an application under development and gather real-time feedback about the behavior of an application. In addition, the system can collect context information that may be useful for the design of an application as well as for understanding user feedback. In order to setup a set of requirements for an ESS, we started by analyzing two context-aware applications that we were developing. These applications were used later to verify whether the ESS that we developed was able to satisfy the initial requirements.

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 648–657, 2011.
© Springer-Verlag Berlin Heidelberg 2011

In the following sections, we introduce related work, design considerations, implementation details, and evaluation results. The last section is for conclusions and future work.

2 Related Work

Experience sampling methods let users to report their experience in situ within their daily life. Unlike a retrospective diary, an ESM directly asks users questions about a certain situation in situ, and, therefore, suffers less from a memory bias problem or a semantic recall problem [10]. An ESM may ask a question like “How you feel about this service?” and let users write an answer on a sheet of paper or type it into an electrical device. An ESM may ask users to report their experience periodically (e.g., hourly), randomly, or driven by an event. Early examples of ESM studies are Larson & Csikszentmihalyi [8] and Brandstaetter [1]. In those days, researchers used a telephone, a pager or an alarm clock to prompt participants to write a diary or answer questions. Since 1990s, the advances of portable devices such as mobile phones and PDAs enabled automated ESM studies [3, 5, 6]. Such devices provided an all-in-one solution of prompting, recording, and retrieving answers.

An experience sampling system (ESS) is for automated ESM studies. Momento [3] is an ESS for evaluating an early prototype of a mobile application in situ. The system consists of a server and multiple clients, and they are communicating by SMS and MMS links. Clients sample pictures, sounds and texts, and a server collects and stores them. The system also supports a Wizard-of-Oz style evaluation method; a server can response in place of a prototype under evaluation. Capturing data is initiated by participants or a researcher. MyExperience [5] is another ESS for evaluating a mobile system. It can record various context data from built-in sensors like a microphone or internal mobile phone events. Captured data is written to a file database in the client. The database on the client can be synchronized with a remote database when the device is connected to an available network. CAES (Context-aware Experience Sampling) [6] is an example that uses context information to control triggering of a self-report. They assumed that people prefer to be interrupted when they are in transition between activities, and implemented a plan-based reminder strategy. The toolkit utilizes an accelerometer and a heart rate sensor to monitor a user and the C4.5 algorithm to detect an activity transition. One of the common problems of an ESS is the intrusiveness of a survey. Txt4l8r [2] proposed an ex-situ diary method supported by multimedia data captured by a mobile device. An ex-situ diary is less intrusive then in-situ experience sampling but may suffer from a memory bias. Txt4l8r tries to help users remember situations better with captured media such as picture and sound.

3 ESS Requirements

We first analyzed the features of existing ESS's with respect to the evaluation needs of the two context-aware applications, iContact and iBell, that we were developing.

Table 1. The features and limitations of existing ESS's

	Features	Limitation / Future work
CAES (2003) [6]	- Chaining questions - Question aggregation - Limitable survey number per day	- Time-driven survey only - Heart rate-based sampling - Attaching GPS, accelerometer
MyExpeirence [5]	- Sensing 140 types of events - XML-formatted survey configuration - Simple context-awareness - Desktop analyzer exists	- No remote access to experimenter - Visualizing data stream - Chaining questions
Momento [3]	- Heavy Server / Thin Client - Using SMS / MMS - Taking picture / recording audio / drawing with pen - Monitoring / Notification support	- User-driven capturing only - Video conferencing - Visualizing data - Reconstruct environment

iContact is an application suggesting a phone number that a user is likely to call in a given context. Two candidate algorithms A & B were being tested for iContact. Algorithm A was based on a conditional probability, for example, the probability of calling someone in the morning on a weekday, which was estimated based on a calling history. The details of algorithm A are described in [9]. Algorithm B was similar to algorithm A but was not based on a conditional probability but was based on a weighted sum of previous call counts in a given context. iBell is an application that suggests a phone's ring mode based on context information such as ambient noise, a user activity, a location, and a history of ring mode changes. iBell's decision is based on "critical contexts" that are predefined situations meaningful for the users. A critical context is inferred by a Bayesian network. When a transition into or out of a critical context is detected, iBell suggests a new ringing mode.

Table 1 summarizes the features and limitations of three existing ESS's ([3, 5, 6]) that we analyzed with the two applications in mind. No existing system satisfied all of our evaluation needs. CAES [6] cannot inspect various activities of mobile phones. MyExperience [5] can perceive rich context of mobile devices, but does not provide an easy way to catch the events of our applications. Moreover, MyExperience's server has limited functionality. For example, its server cannot distribute a research configuration file to clients. Momento [3] aims at evaluation of a very early prototype focusing mainly on qualitative studies.

One of the most desired requirements was that an ESS should have a server-client architecture for scalability and manageability. In order to receive feedback from participants in real-time, we needed a centralized server to monitor and gather data from remote locations. SMS and MMS are handy but may be too limited in bandwidth to deal with frequent and large data transfers. Client devices had better connect to a server via a wireless network such as WIFI and 3G data communication link. An ESS should not interfere with or degrade the original functionality of a mobile phone. We also hoped to minimize users' burden. For instance, users should be able to delay a reply to a survey when they are busy. Finally, extensibility was an important requirement. For example,

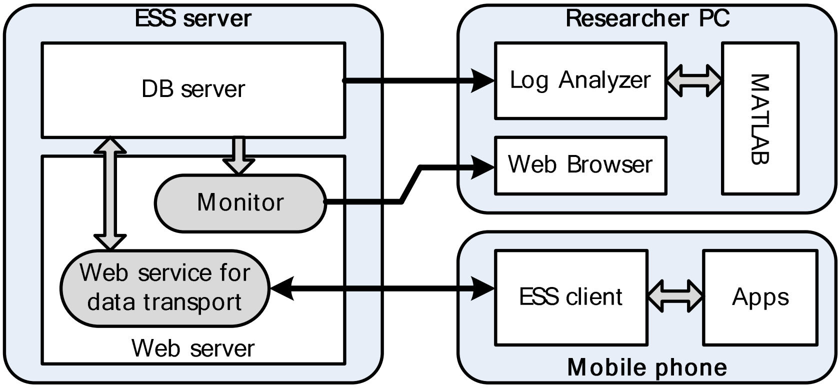


Fig. 1. The outline of the final ESS that we designed based on the ESS requirements

survey questions must be configurable without re-distributing an ESS. Also, the configuration of the system for an experiment should be also easy for researchers.

4 Implementation

Figure 1 shows the outline of the final ESS that we designed based on the ESS requirements. The ESS consists of an ESS server, multiple ESS clients, and researcher PCs. The ESS server consists of two components: a DB server and a Web server. The DB server stores all answers and context data from the clients. The Web server provides an interface to the DB server for the clients and researcher PCs. The researcher PC has two main functions; one is to provide a real-time display of the research status and the other is to provide a tool to analyze and visualize the data from the server. It is required that the ESS clients have a constant connection to the server.

4.1 ESS Clients

The ESS client runs on a Smartphone with Windows Mobile 5. The main function of the client is to gather contextual information and ask survey questions to users. A survey question is inquired when a predefined condition is met or at random moments. A condition is defined by a combination of events and context data. There are 20 detectable events and context date such as making a call, launching an application, and average ambient noise level. Survey conditions are stored in an XML file that is automatically downloaded from the ESS server. The ESS tries to avoid interrupting users when they are busy. For example, the ESS will not ask a question when an important calendar entry is active. The client also records reply times taken for answering a question. If reply time is too short, one may regard that an answer is given without paying attention. On the other hand, a question may have been asked at a difficult time if reply time is too long. An answer given by manual typing by a participant is added to the multiple-choice list for the same question in the next

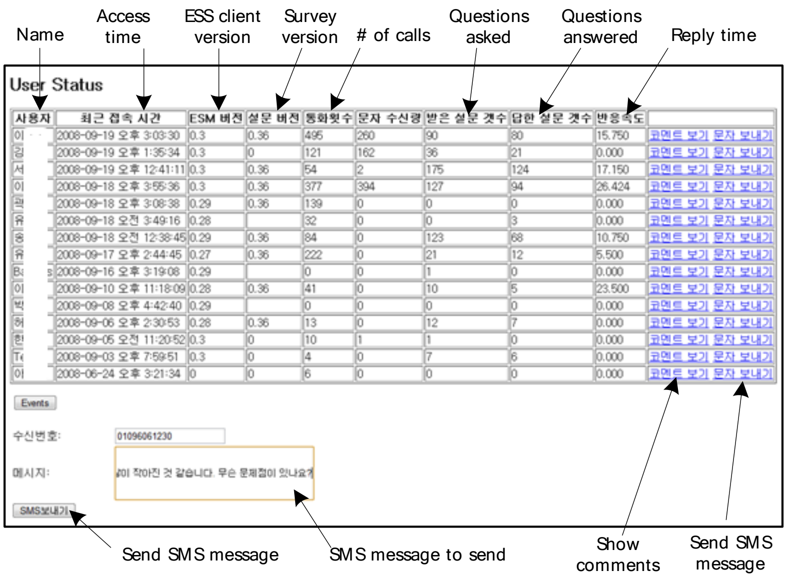


Fig. 2. Monitor: a web page to observe the mobile phone activities of participants in real time

inquiry. A participant can delay answering a question and have another chance later to answer it.

The client consists of the following four modules: Sensor, Core, GUI, and Utility. Each module is designed to have minimum dependency on others for easy modification. Sensor gathers contextual data and events from a mobile device and Core records and analyze them. Core checks whether a survey condition is met and, if it is, asks a question. GUI shows the question which is inquired from the core. The client was designed to be application-independent so that we can extend the system for future applications. In order to add a new application to the ESS for evaluation, a researcher needs only to add some function calls to the application code for communicating with the ESS client.

Some additional functions were added as user experiments went on. One of them is “Diary”. A participant can make a multimedia snapshot of a situation that may be useful in writing a diary later. Pictures, sounds and even videos can be gathered and managed by the ESS server. Comments can be added right after the media is captured, or later through a web page (Diary Logger page) on the ESS server.

4.2 ESS Server

As shown in Fig. 1, the server has a database server to store data from clients and has a web server for researchers and participants to access the data using a web browser. It stores survey configuration files for individual participants and distribute them to clients whenever they are changed. The server also has the function to maintain the ESS software on clients up to date. At first, we tried FTP and socket transfer using

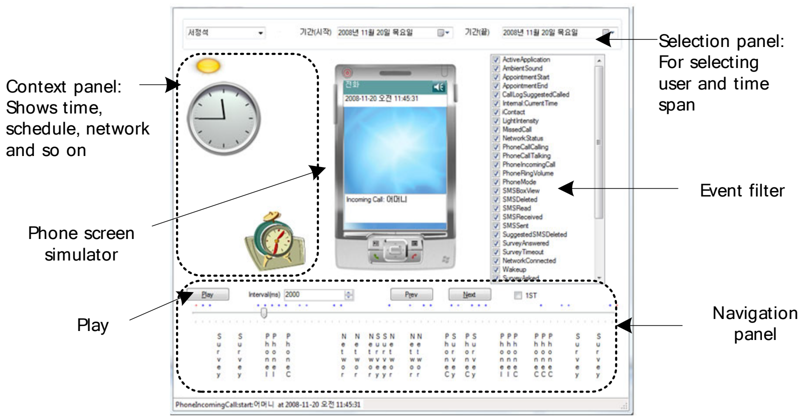


Fig. 3. Context Replayer: a tool for easy browsing and understanding of log data

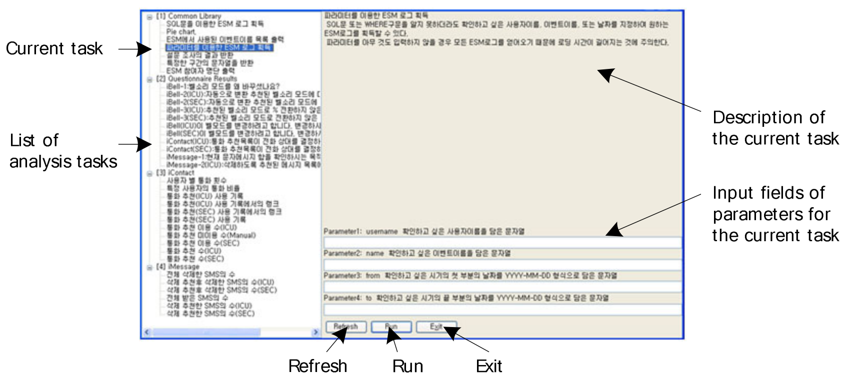


Fig. 4. Log Analyzer: a tool for post analysis of log data

byte stream but experienced firewall and/or NAT problem. We finally settled down with XML-SOAP web service for data transfer between the server and clients.

"Monitor" in Fig. 2 is a web page to observe the mobile phone activities of participants in real time. Monitor is implemented as a web site so that researches can access it from anywhere without a special program. The page shows statistics of a participant, including the latest access time, the survey configuration version, and the number of inquiries so far. The page also has an SMS interface so that researchers can communication with the participants. Participants can also use Monitor to observe the status of the research.

4.3 Tools for Analysis

We made two data analysis tools for researchers to use on their computers. The first was "Context Replayer" shown in Fig. 3. It is a visualization interface for easy browsing and understanding of recorded events and context data. It shows stored events on a time-line as filtered by user-specified options and plays the filtered events like a slideshow. The second was “Log Analyzer” as shown in Fig. 4. It is a tool for the post-analysis of ESS logs. Researchers may want to know the number of times a certain participant answers an inquiry, positive answers about the behavior of an application, the distribution of answers to a specific question and more. Log Analyzer shows a list of analysis tasks and let a researcher choose a task and specify required parameters. Log Analyzer is in fact a graphical interface to MATLAB scripts. We decided to use MATLAB because many researchers already familiar to MATLAB, it provides a good interface to databases and visualization methods, and the functionality of Log Analyzer may be extended easily by adding a new script.

5 Experiments

5.1 Evaluation of Basic Functions

The goal of the first experiment was to check whether the system satisfies the basic functions for evaluating the two applications that we explained in Section 3. Sixteen graduate and undergraduate students participated in the evaluation for 2 to 3 months. All participants were asked to use distributed Smartphones for the experimental period. We paid them for call charge up to about $50 per month.

We gathered 369,974 events from the ESS client saved as 112.609 MB text data in the database. We could perform many kinds of analysis with the logged data. For instance, we studied the usage and satisfaction of two different algorithms A & B for iContact application. The successful suggestion rate of algorithm A was higher than that of algorithm B (27.1% vs. 17.7%), and the satisfaction rate of algorithm A was also higher than that of algorithm B (85% vs. 68%). Also, algorithm A received more positive subjective feedback from the participants. For iBell application, we studied how many iBell suggestions were accepted, and could see that the acceptance rate was below 50%. An inquiry was made after every iBell suggestion to measure its satisfaction rate; 50% answered that an iBell suggestion was useful, 30% answered not useful, and 20% did not answer. The following is a few examples from participants' comment about iBell suggestions: “I don't know the reason for the recommendation”, “Satisfied, but a suggestion is made too often”, “It was already at the minimum volume and the meeting was nearly over”, “Asked repeatedly”.

We could also use the ESS for studying device usage patterns such as call patterns and ring preferences. For ring preferences, melody-mode (37%) and vibration-mode (36%) were the most preferred. We were able to observe how the participants use SMS services. We examined all messages that were deleted as soon as they arrived. 55.8% of these messages were from people in the address book. The average length of these messages was 28 characters. Participants opened these messages 0.3 times on average before they deleted the messages.

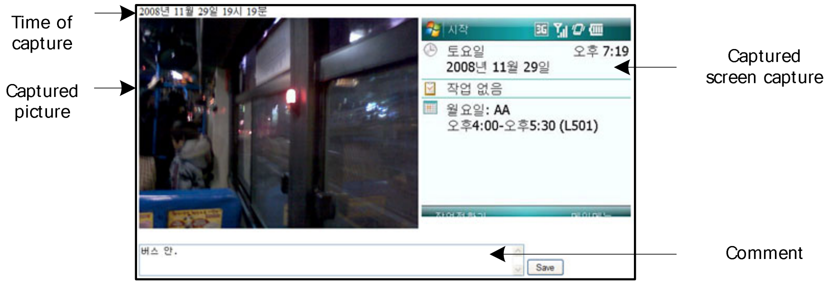


Fig. 5. An example Diary capture shown in Diary Logger

5.2 Long-Term Usage of the ESS

The second experiment was for studying the behavior of participants in a long period. It was done with 7 participants from the first experiment for one month. We could collect 39,675 events stored as 11.0 MB of text and multimedia data in the database including 120 diaries (Diary logs).

With Monitoring Site, we could discover problems in the ESS early and react quickly to them. One participant appeared not accessing the server. We contacted the participant and found that there was a problem in the database. We could notice that another participant was receiving too many survey questions compared with others, and we remedied the problem by changing the participant's survey configuration.

We could infer that participants' motivation was decreasing since we could observe that the comment frequency and the Diary logging frequency was decreased as experiment went on. We conducted a survey to find out the reasons for the motivation drop, and the most frequent answer was “repeated survey” (5 out of 10 answers). We could use the ESS logs to discover reasons behind some user feedback. For instance, to find out what makes participants feel a survey intrusive, we examined logged events before a user answered that a survey is intrusive. The most frequent event was foreground application switch, which means that participants were using an application before a survey was inquired.

The Diary function turned out to be useful. They could describe their experience better, and we could understand their daily life better. For instance, we could observe that one participant manages all schedules with the mobile phone through a Diary snapshot. An example Diary capture is shown in Fig. 5.

We conducted a questionnaire study after the second experiment. Four participants answered that the ESS was inconvenient due to its frequent errors. Interestingly, some of the participants could not distinguish the two applications and the ESS. Errors from the ESS made the participants perceive the applications as unstable. We also inquired about their volition changes during the experiment. One participant answered that the will to participate in the research did not change throughout the experiment while the others answered that it decreased. One of the participants commented that the experiment could have been better if there were more survey questions. However, he pointed out that repeated questions made him bored.

Finally, we asked participants how they used Monitoring Site. Three participants answered that the site was useful. Two participants expressed a concern about privacy. One participant suggested the idea of a communication channel between participants and researchers in the Monitoring Site.

6 Discussion

The Diary function of the ESS was shown to be useful for describing situations that cannot be reported by answering a survey question. Moreover, captured media was helpful for understanding users' life. It is desired to make the function more attractive so that participants may use it more frequently. For instance, users may be more motivated if they can post captured Diary data to their mobile blog at the same time.

It is important to find ways to maintain participants' will to participate in an ESS study. In the ESS experiments, we could observe that many participants showed interests in the ESS itself. One way to motivate participants is to let them engaged in a study more deeply so that they may feel an ownership about a study. This will be possible by letting participants know how a study is going on and share the new findings and conclusions of a study. A communication channel between participants and researchers will be also helpful for promoting their ownership about a study.

7 Conclusion

We designed, implemented and evaluated an ESS for evaluating context aware mobile applications in situ. Throughout the whole cycle of the development, we could identify the requirements and the potentially useful features of an ESS as well as the problems of an ESS that we need to solve in the future. The ESS was still undergoing changes when it was used in the user experiments, and therefore it is still too early to draw any conclusion about the value of the approach and the results. However, we could learn many lessons in the process and could report many valuable findings about an ESS. We firmly believe that our ESS will be able to evolve into a viable tool for evaluating mobile applications as well as studying users in a mobile environment. In immediate future work is to improve the stability of our ESS and incorporate into the system new ideas to maintain participants' interests in an ESS study.

References

1. Brandstaetter, H.: Emotional responses to other persons in everyday life situations. Journal of Personality and Social Psychology 45(4), 871–883 (1983)
2. Brandt, J., Weiss, N., Klemmer, S.: txt 4 l8r: lowering the burden for diary studies under mobile conditions. In: ACM CHI 2007 Extended Abstracts, pp. 2303–2308. ACM Press, New York (2007)
3. Carter, S., Mankoff, J., Heer, J.: Momento: support for situated ubicomp experimentation. In: ACM CHI 2007, p. 134. ACM Press, New York (2007)
4. Consolvo, S., Walker, M.: Using the experience sampling method to evaluate ubicomp applications. Pervasive Computing 2(2), 24–31 (2003)

5. Froehlich, J., Chen, M., Consolvo, S., Harrison, B., Landay, J.: MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In: MobiSys 2007, pp. 57–70. ACM Press, New York (2007)
6. Intille, S.S., Rondoni, J., Kukla, C., Ancona, I., Bao, L.: A context-aware experience sampling tool. In: ACM CHI 2003 Extended Abstracts, p. 972. ACM Press, New York (2003)
7. Kaikkonen, A., Kekaelaeinen, A., Cankar, M., Kallio, T., Kankainen, A., Kaikkonen, A.: Usability testing of mobile applications: a comparison between laboratory and field testing. Journal of Usability Studies 1(1), 4–16 (2005)
8. Larson, R., Csikszentmihalyi, M.: The Experience Sampling Method. New Directions for Methodology of Social & Behavioral Science 15, 41–56 (1983)
9. Lee, S., Seo, J., Lee, G.: An adaptive speed-call list algorithm and its evaluation with ESM. In: ACM CHI 2010, p. 2019 (2010)
10. Scollon, C.N., Kim-prieto, C., Diener, E.: Experience sampling: promises and pitfalls. Dordrecht. Social Indicators Research Series, vol. 39, pp. 157–180. Springer Netherlands, Dordrecht (2009)

Approaching Warp Speed!: Examining the User Experience in the Age of 4G

Quinn Simpson and Sce Pike

Citizen, 2001 NW 19th Avenue, Suite 103
Portland, OR 97209 USA
{quinn,sce}@citizensof.com

Abstract. Using the latest trends of the mobile industry and firsthand knowledge gained from working with a premier mobile network operator, this examination focuses on the development of 4G technology as it relates to user experience. It considers the next digital innovations that this technology may inspire, as well as the future of 4G user experiences. This investigation also discusses how shifting business models could benefit some carriers as they evolve into *smart pipes* and take charge to create amazing user experiences for the next generation.

Keywords: 4G, speed, mobile, apps, applications, network, mobile network operators, MNO, carriers, real-facing, user interface, user experience, cloud computing, mash-up, internet of things, mLearning, Citizen, smart pipe, dumb pipe, cloud brain, pre-learning, post-learning, social reality.

1 Introduction

The world's technology is evolving rapidly, and soon, we'll be hitting the warp speed of 4G connections for our mobiles. But what exactly does it mean to be within the realm of 4G? What foreseeable adjustments will come to mobile applications and user interfaces, and how will this change the way users interact with digital technology?

In short, 4G technology translates into faster download speeds for our digital networks — significantly faster. The International Telecommunications Union (ITU)[1] has set the benchmark for 4th Generation download speeds at 1 Gb per second for stationary connections and 100 Mb per second for mobile [2,3]. When compared to the current speeds of 3G networks, 4G technology will increase download speeds 500-fold for stationary connections and 250-fold for mobile [3].

Many of the top mobile network operators (MNO) in the United States would have consumers believe that they are already operating at such standards, but this is not accurate. While current 3G networks are operating at markedly higher rates than the first 3G systems, network speeds have not met the ITU's standard for data speeds. These carriers promote 4G networks as marketing tools [3], appealing to the average

[1] The International Telecommunications Union (ITU) is an agency of the United Nations providing information and technology solutions to governments and private organizations around the world [1].

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 658–667, 2011.
© Springer-Verlag Berlin Heidelberg 2011

consumer's sense that 4G must be new and improved from whatever device he or she is currently using; after all, 4G is 3G plus one, right?

2 Gauging User Reaction to 4G Speeds

Ironically, as these MNOs promote their misnamed network speeds, they are actually hard at work establishing the infrastructure to handle a true 4G data grid. This means finding a balance between underused fiber optic cable structures and overloaded wireless towers. This means bracing for an evolution of user interface designs and possible revolutionary experiences as our online and offline worlds converge at an even higher rate. But even with advancements to the network and infrastructure, what will 4G speeds mean for users? In this race for 4^{th} generation technology, will users even notice the winners?

2.1 Stage I: First Reactions

When MNOs are finally able to offer authentic 4G connections, it may take several months before consumers see what all the hype was about. In fact, the initial consumer reaction may be no reaction at all, for users cannot take advantage of savvy new infrastructures while designers and developers work through the learning curve of writing to the new platform.

We can look to the iPad for a nice example of this trend. Only after several months of learning Apple's new platform and tinkering with the new specs did developers fully begin to grasp the tablet's capabilities. After the iPad's release, most iPad applications were simply iPhone apps that had been re-tooled to fit the tablet device. There was very little consideration for the tablet's unique characteristics, as most mobile marketers and app developers raced to be among the first to enter the iPad arena. With time, though, developers were able to concept, design, and release apps that explored the new technology more thoroughly and made these user interfaces unique from iPhone presentations. Today's iPad apps are less iPhone-esq and more iPad-unique — and it stands to reason that app development for a 4G network would follow the same path.

The rise in network technology will force a rise in hardware and software technology, requiring new mobile devices and products capable of handling 4G speeds [4]. These devices will expose users to new capabilities and encourage curious consumers to enter the 4G market [5] — all while developers hustle to create slick new user experiences.

2.2 Stage II: Chaos Abounds

Technophiles will be the first adopters of the new devices, and as consumer interest grows, many operators will enter the app market to compete for larger shares. There will be a flood of supply with thousands of apps appearing every day, and yet most if not all of them still will not meet the potential of 4G.

Larger developers and operators will rely on increased mobile marketing budgets to establish brand awareness and brand loyalty among the earliest users. Amid all this clamor, developers will undoubtedly experience an incredible amount of frustration as they battle the challenge of cross-platform fragmentation [6]. Different platforms, different brands, different screen sizes — they all add up to costly app maintenance and an increasingly difficult arena for delivering consistent user experiences.

In a recent TechCrunch interview, Facebook CEO Mark Zuckerberg expresses his frustration for what has already become a problem for 3G app developers:

> *Today it's like, okay, we want to go build an app — even a new product that we launch. We're working on Questions, and it's like, okay, so we build Questions for the Web; then, we build the "m" site for Questions; then, we build the Touch HTML5 version of Questions. Then, we build the iPhone version of Questions, and then the Android version, and then maybe … the iPad stuff. And then, we don't work on a RIM version, and then a bunch of people are pissed because it's not available on their phone. It's kind of a disaster right now. I really hope that the direction that this stuff goes in is one where there's more of a standard.* [7]

With nothing to reconcile this divide among devices, this chaos will continue through the initial launch of 4G networks. It will likely alter the user experience from platform to platform, causing frustration for consumers who prefer standardized experiences.

2.3 Stage III: Unprecedented Market Saturation and Evolved User Interfaces

Chaos in the emerging 4G market will continue to grow as mobile devices achieve higher grades of market saturation. Increased penetration in India and Africa will allow the network to connect over 1 trillion devices [8], and regularly accessing the 4G network will become an ubiquitous experience among the world's citizens. The term *mobile* will refer to more than just phones; it will describe e-readers, portable media devices, new tablets, net books, and more. The amount of digital data generated by users and software producers will grow exponentially as information is sourced from every aspect of the consumer lifestyle.

The amount of data accessible by the network will reach new heights. There will be competition among content developers to create engaging user experiences through new user interfaces. This is where we will likely see the next revolutionary consumer products to impact our day-to-day routines. The more involved user interfaces will take advantage of multiple media types and media sources across multiple platforms. There will be a strong focus on contextualizing presentations as well to create more personalized user experiences.

The influx of data and the sheer size of the network will drive the development of cloud computing. This innovation could solve the issue of fragmentation across platforms as each device pulls the same data directly from the cloud. Designers will be able to deliver consistent user experiences within a much more reasonable framework for development.

3 The Day after Tomorrow

Based on the forecast for achieving a true 4G network, here are a few trends we can now predict.

3.1 M2M Communications Create an Internet of Things

As mentioned, cloud computing will become commonplace within digital industries as machine-to-machine (M2M) communications reach a new level. Over 1 trillion networked devices all over the world [8] functioning at 4G speeds will transform the Web into an Internet of Things, and when combined with a smart grid power structure, the network will offer consumers new ways of controlling their everyday lifestyles. Personal mobile devices will automatically adjust their users' home environments, appliances, vehicles, and more through contextual awareness and emotional recognition. Mobile broadband will also allow all users to engage with the Internet of Things from any place at anytime [4]. Every smart device will produce data to go along with new software and information generated by users.

3.2 Mash-Up Interfaces

With such a massive amount of data to consider, mobile app developers will design programs to shape interactive, adaptable, and predictive digital experiences. These will be achieved through mash-up interfaces, which will present confluences of data as individual, interpretable experiences. These customized environments will represent supreme destinations tailored to respond to users' personal preferences, emotions, moods, and physical locations. Apps will contextualize interactions, creating more sensory experiences that engage and respond to user interests.

For a primitive example of what a true mash-up environment will deliver, we can look at today's version of Google Maps. Through a single interface, users can view any physical location on the planet, expanding out to see it from a cosmic perspective or drilling down to see the space on a local level. Users can also take their explorations a step further by viewing the location from the street itself. For even more information, users can toggle layers to appear on top of the map, showing traffic flows, user-submitted photos of the area, terrain estimations, and live feeds from area webcams.

With 4G, these types of interfaces will exhibit an unprecedented amount of software contextualization and adaptability. Reality as well as individual user experiences will add so much more data to the system (and in real-time) so that the user interface will present life as it is, enhanced. There will be less reliance on creating virtual worlds to communicate lifelike experiences; instead, the 4G network will simply tap reality as the basis for these experiences.

3.3 *Real-Facing* as Mixed Reality

This convergence of online and offline lifestyles will bring us closer to mixed reality presentations in which video, audio, and 3D iterations merge real data and virtual data to create augmented reality [9]. For user experience designers, this means *real-facing* with users, or in other words, designing user experiences that are so predictive, so adaptable and intuitive, they approach the users' realities.

Over time, as designers push the capabilities of the new system, the ease of use for consumers will increase dramatically. New real-facing interfaces could emerge to overcome multiple barriers between the user and the interaction. Similarly to how the iPhone's touch screen capabilities allow users to interact with programs without relying on a keyboard and mouse, new devices will materialize to reduce barriers even further. And, as this trend continues, it may become difficult for consumers to delineate between reality and the virtual augmentations provided by various programs.

As programs reach this level of efficacy, not only will they incorporate new ways of connecting users, they will also introduce more and more behavioral dynamics to solicit, encourage and provoke specific user actions [10]. With the line between the virtual and real worlds blurring even further, barely visible layers of governing dynamics may be applied to programs to create wholly immersive user experiences.

Already we see programs utilizing gaming dynamics to influence user behaviors in this way. Games like Foursquare, Farmville, and The World of Warcraft are able to coerce a variety of human responses, including appearing in specific physical locations, logging in to play at specific times, and working through a series of tasks to achieve specific rankings. These same dynamics can be added to current systems to enhance their effectiveness, to create better frameworks, to incentivize specific responses, and to focus user attention into collective efforts [10].

3.4 mLearning

mLearning represents arguably the biggest departure — technologically speaking — from how consumers function today. The concept begins with 4G mobile devices integrated into classroom and work settings, streaming multiple video and audio presentations simultaneously, making every textbook, professional presentation, or creative example accessible from the cloud.

Some of these innovations are already taking root. In January of 2011, Amazon announced that its e-book downloads have eclipsed its sales for paperbacks for the first time in history [11]. Meanwhile, schools around the United States are already requiring students to acquire mobile devices in an attempt to eliminate bulky, hard-bound books [12,13]. Progress like this may make it seem like mLearning is here to stay, done and done — but, let’s take learning outside of the classroom for a moment. Let’s talk about the way we learn in general and to what purpose.

Traditionally, the national school system has functioned to educate our citizens so that they can be better prepared to enter the workforce. This means students spending years in classrooms absorbing concepts over a range of subjects so they can become functional employees, capable of recalling and processing information when their professional roles call for it. In other words, individuals depend on a certain amount of *pre-learning* to help perform tasks.

One trend we are already noticing that may eventually challenge this dynamic is the speed at which individuals are able to acquire information online. Through this increased speed, we anticipate seeing more of the phenomenon known as *post-learning*, which takes place when an individual relies on information gathered *after* an assignment is given in order to perform the integral processes needed to complete the task. As opposed to relying on a formal education to supply the arguably limited knowledge of that particular concept, the individual acquires the knowledge as needed online. The individual can rely more heavily on post-learning and be able to focus pre-learning efforts more effectively as a result.

3.5 Your Cloud Brain

Only with 4G speeds and beyond will this idea of post-learning evolve into a practice that leads us all to question the necessity of traditional classroom dynamics. With so much digital data accessible so easily, a social preference for pre-learning may shift to a preference for post-learning, allowing the cloud to produce answers more frequently than our own internal knowledge.

The idea of human beings relying more heavily on automated systems and machines for influencing day-to-day interactions may make some social observers nervous. But, researchers see more promise in this path than doom. In a recent TED Talk, anthropologist Amber Case contends that technology is not taking over but really fading into the background, enhancing, allowing us to connect more and overcome even more barriers [14]. In other words, the future will not trap us in some virtual world controlled by machines. Instead, it will deliver reality to the n^{th} degree — the enhanced reality described earlier in which technology holds an ambient presence in our everyday lifestyles.

Faster speeds allow individuals to be connected at all times, to have instantaneous access to information, and to rely more on the cloud for personal data storage. This is very different from cloud computing, as the individual's personal data, life story, and memories are all documented automatically in the cloud. This evolution will relieve individuals from the clutter of "menial" memories, like where they parked their cars, and allow them to focus their brain power on more meaningful thoughts. In essence, the network becomes the place for individuals to find information that may not be necessary to carry with them at all times. It functions as the individual's second brain: their Cloud Brain.

4 How 4G Speeds Will Affect MNOs

As the race for being among the first networks to achieve 4G technology continues, mobile network operators are positioning themselves to lead the next evolutions and revolutions in digital industries. But what exactly will a 4G network mean for these carriers? How will the innovations of today shape the innovations of tomorrow?

Currently in most markets, MNOs are experiencing their highest streams of revenue through SMS/text message services [5]. But, with the infusion of faster data transmittance, carriers could see a change in their business models. Even though rates for content and applications represent only a small percentage of MNO revenue today, they may represent the biggest opportunities for future growth [5]. The adoption of next generation devices by consumers will help fuel this increase in revenue, but the real high-octane impetus for new earnings will depend on the MNO's willingness to be more than just a network provider — and to instead become a provider of services. It is a question of continuing business as a *dumb pipe* and simply transmitting data to users, or evolving into a *smart pipe* and having a hand in delivering the most innovative user experiences.

5 The Speed of Culture

One of the many things that we learned with the release of the iPhone is that user experience matters. Captivating, perceptive interfaces will attract a large number of consumers. This is why some of the major MNOs have partnered with creative agencies to develop anticipatory 4G software; they see the opportunity to cultivate brand awareness and to plant the seeds that will likely yield high returns.

Our company, Citizen, is one of these agencies, and we are working tirelessly to usher in the next generation of mobile applications. We are currently working with a major U.S. mobile network operator to provide cutting-edge 4G user experiences. This MNO is one of the few American brands with the infrastructure, capacity, and large customer base to develop as a smart pipe with an established following. At Citizen, we like to say that we operate at "The Speed of Culture," and in examining our latest work with this MNO, the phrase seems appropriate.

To generate state-of-the-art user experiences, we dig deep to understand our client brands. We want to get to know cultures, customers, and how they operate. We will actually send our people to work on location with our clients in order to absorb their brands. For this particular MNO, we partnered with the company to become more of a team member than a consultant, and that relationship has allowed our projects to accelerate quickly.

We formed three apps as part of an initial assignment to guide this MNO into the 4G age. The concepts for these apps focus on some of the existing techniques for 3G app development and enhance them. They play on the most obvious advantage of developing for 4G speeds: being able to display massive amounts of data at once. Specifically, the applications explore multiple live streaming broadcasts of video and the live social interaction around them [Fig.1,2]. They also explore the use of crowd-sourced content to further enhance and contextualize user experiences. Essentially, we took the now simple integration of video and brought it to the 4G level. This was new; this was fun. But, our next set of apps really started to push our creativity.

With our second round of work, we ramped up our thinking to introduce bits of reality into the overall concepts. Our first application defines an interface around users' relationships with their wireless carriers, using visuals to show their usage levels, billing information, applications, and online visibility [Fig.3]. We designed an interface that presents this data as a virtual environment, complete with bodies of land that represent suites of applications, as well as cloud density to show cloud-based data and online visibility. We took this idea a step further and designed a game around the environment. Increased wireless usage triggers new creations in the environment, incentivizing users with new experiences and encouraging them to explore their wireless applications even more.

Our second application explores the gaming dynamic further, with even stronger reality-as-interface cues and with a social twist. The game is driven by users taking photos of their environments and sharing them with friends in a competitive manner. In essence, this work incorporates the user's real-world environment into the software's functionality to present a *social reality*, where the social network is closer to the user's real world than what we see today with Facebook or Twitter. It is the real world augmented with cloud data.

We continued with this line of thinking to develop our most recent work, which we believe delivers legitimate, contextualized 4G experiences. These apps use mash-ups of all available data to recreate the real world in a sense. One application is a navigation tool that uses the device's viewfinder as the primary interface and enhances it with a wealth of contextually relevant information [Fig.4]. Another application turns a maximum capacity stadium into a sports fan's dream by revealing spectator tweets, related media, locations, thoughts and predictions above the heads of their originators. The primary interfaces for these apps *are* the users' realities. The apps then enhance the user experiences by incorporating all types of data from the Web, from social environments, and from the sensory hardware in the device itself. They predict; they adapt; they react. These are the most cutting-edge applications we've developed to date in anticipation of 4G speeds. There is no virtual environment presented to the user at all; instead, reality serves as the interface.

Fig. 1. Multiple live streaming broadcasts of video and the live social interaction around them

Fig. 2. Multiple live streaming broadcasts of video and the live social interaction around them

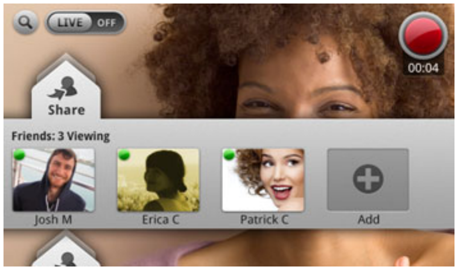


Fig. 3. An interface around users' relationships with their wireless carriers

Fig. 4. Navigation tool that uses the device's viewfinder as the primary interface and enhances it with a wealth of contextually relevant information

Recognizing the innovation that 4G speeds will demand has allowed our client the opportunity to be a leader in the smart pipe arena. In this regard, we feel very privileged to be working with a carrier that recognizes the importance of pushing the creative envelope. It helps us both stay a step ahead of the game — and that's important for when the speeds really start to pick up.

References

1. International Telecommunication Union, http://www.itu.int
2. McPeek, M.: What Does 4G Mean to Me? TopTenReviews. TechMediaNetwork. Online article, http://cell-phone-providers-review.toptenreviews.com
3. Ziegler, C.: 2G, 3G, 4G, and Everything in Between: An Engadget Wireless Primer. Engadget. AOL. Online article, http://www.engadget.com
4. 4G Applications Set to Evolve in a Multiplatform Environment. Maravedis. Online article, http://www.wimax-industry.com/ar/26m.htm
5. 3G and 4G Mobile Services: Is It the Application or the Device? Communications Infrastructure and Convergence, vol. 3(2). Stratecast, Frost & Sullivan (2009)
6. Husson, T., Ask, J.: 2011 Mobile Trends. Forrester Research. Market research report (2011)
7. Arrington, M.: Interview with Mark Zuckerberg on the "Facebook Phone." TechCrunch. AOL, http://techcrunch.com

8. The Future of Mobile Media and Communication. In: Mocom 2020. Online video presentation, http://wn.com
9. The Mixed Reality Lab (MRL), University of Nottingham, http://www.mrl.nott.ac.uk
10. Priebatsch, S.: The Game Layer on Top of the World. In: TED Conferences, http://www.ted.com
11. Amazon Kindle E-book Downloads Outsell Paperbacks. Online newspaper article, http://www.bbc.co.uk/news/business-12305015
12. Collette, C.: Minnesota School Replacing Textbooks with iPad. WSTP.com. Pacific and Southern Company. Online newspaper article, http://www.wtsp.com
13. Webb School of Knoxville, http://www.webbschool.org/technology
14. Case, A.: We Are All Cyborgs Now. In: TED Conferences., http://www.ted.com

Methodology for Evaluating Multimodal Biometric Authentication on Mobile Devices

Claudia de Andrade Tambascia, Robson Eudes Duarte,
and Ewerton Martins Menezes

Telecommunications Research and Development Center
Campinas, São Paulo, Brazil
{claudiat,robsond,emenezes}@cpqd.com.br

Abstract. Biometric authentication systems are an important component in more sensitive applications, since, in conjunction with the use of a mobile device, enable so-called "multifactor authentication", which involves something the user knows (a secret, e.g. password or PIN), something he has (mobile device) and something that he is (a biometric trait). This paper aims to present usability recommendations to be followed during the implementation of applications or services for mobile devices. These recommendations will be presented taking into account international standards and standards already used in other applications and devices. As a result it is expected a set of specific recommendations for the technology that will be adopted in the context of BIOMODAL project.

Keywords: Biometric authentication systems, Usability recommendations for mobile devices, Multifactor authentication.

1 Introduction

When thinking about mobile devices and their applications, it becomes possible to say that today the mobility aspects are increasingly common and necessary for society as a whole. Users of this type of technology have a different usage experience of desktop and notebook users, which demands that computer systems developed for mobile technology take into account some specific features to ensure that usability requirements are met. When these requirements are not met, it is often possible to associate users non satisfaction with aspects related to:

- Hardware devices with a design non appropriated for use;
- Information and content with none customization to the mobile environment;
- A lot of information presented on the screen, making the location desired by the user's option very difficult;
- Content loading time inappropriate, due to slow connection speeds, stopping interactions [1].

The purpose of this paper is to present usability recommendations that should be followed during applications and services development for mobile devices, and contributing to ensure the quality of experience of end-users. These recommendations

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 668–677, 2011.
© Springer-Verlag Berlin Heidelberg 2011

are based on international norms and standards and will be adapted according to mobile technology that is assumed and the prototype which will be developed in the context of BIOMODAL project, thus allowing the development of a methodology for usability evaluation to be applied on these applications and services.

2 Definitions

The term Usability, defined in [2], is related to the measure by which a product can be used by specified users to achieve specific goals with effectiveness, efficiency and satisfaction in a context of use. Among the specific objectives to be measured and achieved are considered aspects that allow the user to reach his/her goal, taking into account the necessary effort to ensure that this goal is met.

Under the BIOMODAL project it was defined, in the first time, the applications development for Android platform, because is an open source and it's growing in the global market. These technologies minimize testing costs and requirements definition that must be followed in the applications development to be available.

In the following section it will be presented usability recommendations used in this specific project for mobile devices, as well as standards considered for this technology. Such information will be used during the creation of Mobile Devices Recommendations Guide (MDRG).

3 State of Art

Despite the BIOMODAL project considers the use of mobile devices like iPhone and Android operating system, usability recommendations were searched for mobile applications embracing also other technologies and it was found that many of these recommendations can be considered universal, and others may be completely dependent on application-specific requirements context. This paper will present some recommendations identified for different devices and recommendations which should be fully considered in applications development for BIOMODAL project.

3.1 Applications for iPhone

Recommendations for iPhone were defined in a guide structured into two parts. The first one describes the iPhone OS environment and types of software that can be developed, covering also fundamentals of interface design and describing how to apply these principles in applications iPhone design [3]. The second part describes several interfaces elements available in the iPhone SDK and how to use them when building an application.

An important aspect outlined in this guide is related much more to the differences between developing applications for mobile devices and computers and how it impacts on design decisions. These differences are related to the small size of the screen, limited memory, and visualization of one screen at a time and, in the iPhone specific case, running only one application at a time. In terms of implementation, native applications must appreciate the ease, simplicity, responsiveness and esthetic interface appearance. For this reason it was defined styles provided by the application based on visual characteristics and behavioral data model and user experience where productivity applications must keep the user focused on the task and information, usually by putting the information in a hierarchical manner. It can also offer customization options, utility applications should require minimal user input, be visually attractive and flexible for different types of content on the other hand immersion applications should occupy the full screen and offer a customized interface, with a visually rich environment focused on content and user experience, being used mostly for fun.

In the BIOMODAL project context development will be held within productivity applications mostly though it will also consider utility applications. Another important aspect to be considered relates to porting applications from desktop to mobile platform. In this case, it is important to consider the different expectations and context of use by the user of the application, because often the use of the iPhone is done in environments full of distractions, with brief task execution and does not require total attention and may be a good practice to focus on the application iPhone in features that cover the greatest number of users.

The features to consider on an interface project for iPhone are elements like metaphors, direct manipulation, choice, feedback mechanisms, user control, aesthetic integrity, intelligibility, organization, text input, informational texts, primary tasks, targets, and effective communication, and for each one, respective recommendations were defined and some of them will be used in BIOMODAL project.

The iPhone recommendations presents a constant concern in maintaining the native operating system standards in applications developed for this environment, such as color scheme, interface elements and their behaviors so the user could transfer the knowledge acquired with the iPhone OS interface for its applications. This might be a recommendation to be followed during the BIOMODAL project.

3.2 Applications for BlackBerry Smartphones

A recommendation guide defined to applications developers for BlackBerry smartphone platform gathers specific tips and good practice design interfaces [4] with great importance for BIOMODAL project.

Applications developed for BlackBerry platform must provide a balance between the best user experience possible and battery consumption level. Taking into consideration the main differences between mobile devices and traditional computers, it must be also listed factors as small screens, processor with lower speed, low memory, battery charging time and the limitation of only one screen be displayed at a time as limiting factors in the development of applications for this type of device.

How the user interacts with the mobile application is also different than the way of interaction on computers applications. In this case, the efficiency and speed with which the user performs the tasks in this environment are only reached understanding user navigation model and providing mechanisms for this kind of interaction happens as quickly and easily as possible.

The features to be considered for this item on a interface project for Smartphones include elements such as navigation, interface components, focus on the task, menu actions, undo function, information presentation, touch screen, keyboard, cursor, application screens , banners, dialog boxes, progress indicators, information and images, menus, fonts, conventions of writing, icons, sounds, location, design of the interface and colors, with their respective recommendations and were also considered in BIOMODAL project too.

3.3 Applications for Android

Developing applications for Android is supported by a dedicated portal where developers found articles about operating system compatibility with different devices and design recommendations focused on performance, responsiveness and treatment of interruptions.

As the Android operating system runs on a variety of devices, compatibility is a very important requirement in their applications. One of the ways to promote compatibility is filtering applications to be installed on a given device according to the hardware characteristics, such as gifts, compass, accelerometer, camera, touchscreen, among others. In addition, another point to be considered on how to increase compatibility is to support the screens of different sizes and resolutions. It is possible, in a file called manifest of each application, set attributes on which sizes and resolutions are supported, where the screens are categorized into large, normal, or small relative to its size, density and high, medium or low.

The features to be considered on a project interface for the Android operating system take into account elements like application screen, presentation, interruption of activity, back key, home key, menus and commands, with their respective recommendations.

Due to restrictions on mobile devices, this type of applications should have a particular concern in relation to the performance, responsiveness and disruptions suffered by the user. In Best Practices is given a detailed approach on how to achieve these goals, in order to keep its code always efficient. It is important to note that the application should not run tasks that the user didn't need nor allocate memory when you can avoid.

3.4 Applications for Other Mobile Devices

Studies conducted by researchers at Americans in [5] made comparisons of limitations and features of mobile devices interfaces developed with desktop interfaces, using the existing interface recommendations for desktop as a starting point for the formulation of mobile devices specific recommendations, presented as rules that are used as the basis of that study and among the eight initial recommendations. Four of them are kept fully and the remaining is adapted to the mobile context [6]. The key features for the mobile device interface design are elements such as shortcuts, feedback, dialogues, sequential actions, control, consistency, rollback actions, errors, memorization, multiple and dynamic contexts, size of devices, modes of interaction, storage and retrieval, interaction top-down, customization, and aesthetics, with their respective recommendations.

From all the studies conducted, researchers have made important discoveries of usability, among which we can mention that in mobile applications the tasks performed by the users should be appropriate to the time and cognitive resources employed in its realization. And the design must be done so that the user can transfer the knowledge acquired with the applications that he/her has already touched.

Due to limited data bandwidth, content intended for mobile devices must be smaller than the normal site content and must not require very large downloads, since losses are very frequent. While vertical screen scrolling is not wanted, scrolling the horizontal should be avoided at all cost because it generates serious flaws in mobile functionality.

3.5 Usability and Biometrics

Aspects related to ensuring successful biometric systems are connected directly with usability-related factors, such as the National Institute of Standards and Technology (NIST). The documentation provided is intended for designers and developers of biometric technologies, describing the process of user centered development and how to incorporate this process can improve the effectiveness, efficiency and user satisfaction in biometric systems.

Currently aspects of users (age, gender, experience, abilities and limitations) have been disregarded in design of biometric systems, with main focus on performance, functionality, reliability and accuracy of system, setting aside the user and his/her influence on the system. In fact, the user is a key piece on biometric process because only with user's acceptance any biometric technology can evolve.

The features to be considered on this item for a project on biometric applications are elements such as effectiveness, efficiency, customer satisfaction, ease of learning and ease of memorization, with their respective recommendations.

4 Aspects Related to the Biometric Systems Development

Although usability and accessibility recommendations are listed in the context of development of biometric systems, to improve the ability of end-users in the use of a product effectively and efficiently, the process of user centered design must be considered to be of great importance.

This process is considered interactive with a lifecycle that is composed of four main phases and which are known to cause measurable impacts on usability and ease of use of biometric systems.

The four main stages of the process are: definition of the usage context, determining the target users and organizational requirements, development of design solution and assessments implementation. Figure 1 illustrates the defined phases.

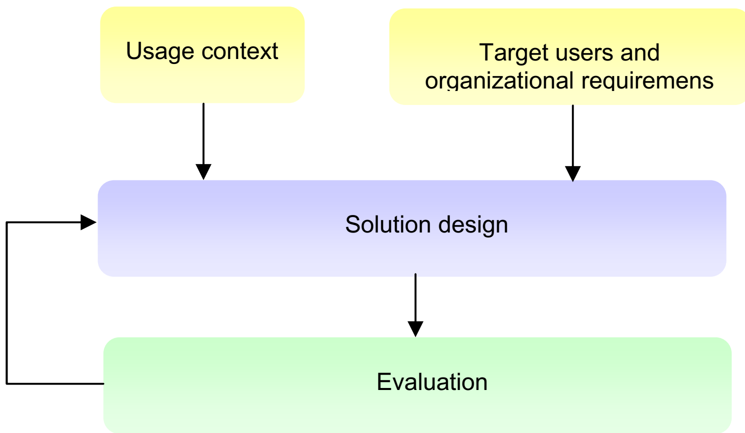


Fig. 1. Phases for biometric systems development

The definition of usage context involves not only the user's context, but also the business environment in which the biometric system is developed, the operating environment where the system will be used and the social environment where it will be implemented. Understanding and sorting system users is an important task in this definition, because it facilitates the development of a system that meets the needs and expectations of users. One way to achieve this goal is to define characteristics of the end users and how they can influence the final design of the system. These characteristics are summarized in Table 1.

Table 1. End user features

Features	Questions to consider
Age	Older users need larger texts or sounds higher tips. Younger users are not necessarily more successful
Genre	Genre may or may not be more successful than another.
Anthropomorphic	User height can affect his/her ability to use the system.
Ethnicity, nationality, language and culture	Without a common language can be no impact on the design of feedback and instruction guide.
Education	Education can affect system usage. Biometric systems should cover users with little or no ability to read.
Experience and knowledge level	Novice user needs differ from the frequent or regular user needs.
Accessibility attributes	Design requirements should be considered to make the system accessible universally.

In addition to the aspects related to the user, environmental factors may influence or produce any effect that can directly influence the design of biometric systems. These factors are described in reference Table 2.

Table 2. Environmental factors influence in biometric systems

Features	Potential impact
Location	Physical environment in which the biometric system is affects or influences the design.
Device location	Device's location affects the user's ability to use or access the system.
Temperature and humidity	External environment affects the system performance and temperature and humidity affects the capture of a high-quality sample.
Ethnicity, nationality, language and culture	Without a common language there can be no impact on the design of feedback and instruction guide.
Lighting	The light level lighting affects the reading or the visibility of the graphics device.
Noise	The noise level can affect the ability of an individual to hear tips and information for audio and give a feedback to the system.
Instructions	If there are instructions, icons and sizes of the letters must be seen as appropriate and understandable.
Help	In a given environment, it must be known the better way to provide feedback, errors and help information.

With the definition of use context it is also defined tasks the user will do in the system, seeking to prioritize them according to user needs, usage frequency, importance, vulnerability, and if it is doable.

In the solution design phase, the design considerations should not be limited to end users, but should consider the users who will operate the system and analyze the data. The goal of the interaction design is shaping the behavior of the system, without pegging it to an interface, so that system and user can work in partnership to achieve the end goal of the application. For each user action that the system should provide a feedback and to make that possible, a solid design specification is required.

As soon as "what to" and "when to" do are defined in interface design, the main point is how to make the communication between the system and the user. A consensus is that the system should provide feedback in a multitude of ways, just in case something fails or the user does not understand the method chosen. In addition to considering the contents, the interface design process must also consider the most effective way to display it to the user.

Finally the last phase is the evaluation process, an essential part of the process because it provides assurance that the design is being developed correctly, assisting in the identification of problems that still need to be resolved.

Evaluation and design are phases that must be done interactively, where the results obtained with the end users in reviews are incorporated into the design making the system meet as much as possible the needs of users. In the initial stages of design, the assessment should be more qualitative, where impressions and reactions of users are of great importance. In the later stages of development, though, the system must be tested by means of more quantitative usability testing, where the performance when using the system is measured.

5 Guide Recommendations for Mobile Devices in the Context of BIOMODAL Project

Based on research described in section 2 of this paper it's possible to determine usability recommendations that should be considered in the biometric recognition prototype to be developed in BIOMODAL project context.

For this purpose it was defined categories where every aspect to consider is detailed, showing its definition and the associated recommendation, looking to illustrate with possible solutions for development.

This process of categorization and descriptions of the recommendations is done by setting the requirements of the application being developed. Therefore, it is necessary to wait for completion of the project requirements document of BIOMODAL for specific recommendations to guide the implementation. What can be done initially is the division of components that will be treated in the interface and interface elements that will be associated and which recommendations should be followed in the development of the prototype. Table 3 presents the elements which should in principle be treated and their groupings, according to the component under consideration.

Table 3. Components and elements to be considered in the BIOMODAL prototype

Element	Components
Window structure	Title bar, application area, scroll bar and secondary window.
Menu structure	Menu bar, number of options, drop-down menu and contextual menu.
Forms structure	Structure check, grouping information, separator lines, input and output groups and commands group.
Information structure	Information group and encoded information.
Driving mechanisms	Guidance, feedback, status, control, error prevention and error correction mechanisms, error message and help engine.
Presentation	Used colors, font style, font size, icons and selection or pointing area.
Interaction	Browsing interaction model, callback option and steps to run the task.

The components and elements defined, associated with its recommendations, will be also used in methodology of validation of applications for mobile devices, and so the recommendations must be clear and well defined, so that developers won't have questions about how to deploy the application.

6 Conclusion

Usability recommendations are considered an important requirement in the BIOMODAL project context to guide developers the implementation of services or applications, ensuring the user-friendliness and satisfactory execution of tasks, contributing to the quality of user experience.

For adequate preparation of the Recommendation Guide, it was necessary to determination the family of devices to be covered by the project, as well as the requirements of prototype being developed, so that in accordance with the Guide, is designed a usability evaluation methodology for applications on mobile devices.

Any system that intends to reach the mobile market has to deal with specific usability issues from the mobile context like small screens, inappropriate design of hardware, problematic input method, inadequate content loading time and the dynamic context of use. We aim mainly to define usability aspects which must be tracked throughout the services and applications development for mobile devices as well as the contribution of these aspects to user's Quality of Experience (QoE) enhancement.

Those usability issues emerge from international standards in the scope of an R&D project of multimodal biometric authentication for banking services on mobile devices. The methodology proposed should determine how to measure, categorize, identify and observe user's interaction with such services in a usability laboratory. Taking into account that the user’s perceptions and expectations of the system are crucial in biometric authentication, it is important not just to improve its performance and effectiveness but also to make it useful in the real world.

To properly run the usability tests defined on the methodology we recommend a lab composed by two distinct rooms. One of them is the test or interaction room, used to host, accommodate and conduct the users who will perform the tasks, equipped with computers, mobiles devices, monitoring cameras and touchscreen monitors. The other is the observation room, an annex building to the test room separated from it by a two-way mirror, used to accommodate the professional team who will follow the tasks execution and manage the tests recording. In order to better organize the usability activities and reach the project HCI goals, the methodology can be customized to the defined test purpose and involves 4 main stages: contextual analysis, test setting, test execution and results analysis and interpretation.

On the contextual analysis the test techniques and the application platform are defined with a range of available options, such as, quali-quantitative evaluations, use of real devices or emulators, simultaneous or consecutive verbalizations and previous heuristic evaluations. The specialists should align the chosen techniques according to the test purposes and platform and for this reason a broad knowledge about the project and its current development status is required.

On the test setting stage it is necessary to determine the number of target users that will perform the test, the tasks which will be included in the test process and define more than one data storage and noting techniques. Moreover it is important to set the test environment so that it would get as near as possible to the real context of use. It is also on this stage that the participants recruiting process happens.

On the test execution stage, a pilot test must take place to ensure that all the previous settings are utterly and properly working. During the test process, it is preferable that only one from the usability team addresses personally to the user to clarify the tasks to be performed, lead and help in case of critical doubt and to start and finish the test. By the beginning of the process, the conductor must explain the test purposes, making it clear that it is not the user who will be evaluated but the developed solution. Annotations must refer the time of relevant events (doubts, confusion, errors) and profile or satisfaction interviews may take place before the test ends.

Finally, the stage of results analysis and interpretation must allow the whole usability team to review the recordings and recover relevant data which might confirm or disprove formerly written hypothesis. In this stage a report must be made, containing the type of test executed, the place, the registration method, the number of target users considered, the scripts and environment, the data collected, the results obtained, list of usability issues identified, including category and priority of the problems found.

References

1. Ballard, B.: Designing the Mobile User Experience. Little Springs Design, Inc., USA. John Wiley & Sons, Ltd, West Sussex (2007)
2. Tambascia, C., Menezes, E.M.: Aspectos relacionados à Usabilidade. PD.30.13.22A.0001A-RT19-AA. Technical report (2010)
3. Apple Inc. iPhone Human Interface Guidelines. Human Interface Guidelines. All rights reserved (2010)
4. BlackBerry Smartphones. Version: 2.3. UI Guidelines (2009)
5. Gong, J., Tarasewich, P.: Guidelines for handheld mobile device interface design. College of Computer and Information Science, Northeastern University, Boston (2004)
6. Shneiderman, B.: Designing the User Interface - Strategies for Effective Human-Computer Interaction. Addison-Wesley, Reading (1998)

Phone Use and Aging: Do Spatial Interface Metaphors Help?

Mark Turner and Roxane Cortes

Department of Psychology, University of Portsmouth, King Henry Building,
King Henry I Street, Portsmouth, PO1 2DY, UK
mark.turner@port.ac.uk,
roxane.cortes@myport.ac.uk

Abstract. This study compared the usability of two different speech activated automated telephone services with a group of older adults, aged between 50 and 87. All participants completed a series of information search tasks with a standard numbered menu system, and with a specially developed menu system designed around a spatial metaphor. Automated menu systems presented a significant barrier for older adults with only 59% of participants able to complete all search tasks. Task completion improved when using the spatial metaphor-based version of the service. The spatial ability of participants was related to successful task completion with both services. Time taken to complete search tasks correlated with working memory capacity only in the case of the standard service and not the metaphor-based service. The incorporation of spatial metaphor within automated telephone systems may therefore provide an important aid to navigation when developing future services with older adults in mind.

Keywords: Aging, Interface Design, Automated Telephones, Spatial Metaphor.

1 Introduction

The proportion of older people in the UK population is growing rapidly. Population forecasts indicate that by 2033, 29% of the UK's population will be over 60 years old (Office for National Statistics, [1]). Reductions in personal mobility and autonomy that occur with age also mean that older adults are becoming more reliant on technology for communication and continued connection with society.

The use of devices such as the mobile telephone and voice activated menu systems now form an essential part of communication within modern society in a wide variety of contexts such as banking, transport and health care; yet their use may present a considerable barrier to independent living for older adults. Older people are less likely to adopt new technologies (Renaud & Van Biljon, [2]), perform less well when navigating through phone-based menu systems (Ziefle & Bay, [3]) and experience more difficulties understanding synthetic speech (Roring, Hines & Charness, [4]).

Age related declines in cognitive abilities are also well documented; older adults have been repeatedly shown to perform worse than younger adults on laboratory tasks

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 678–686, 2011.
© Springer-Verlag Berlin Heidelberg 2011

involving memory and controlled temporal processing (e.g Kausler, Salthouse, & Saults [5]); such skills are commonly required when interacting with serially presented auditory information via telephone. A more effective means of presenting information through automated telephone menu systems is therefore required to support the domestic needs and continued inclusion of older adults in society.

Previous research by the applicant with young adults has demonstrated that voice-activated telephone systems based around the use of a spatial interface metaphor are effective at improving caller performance compared to standard telephone menu systems (Howell, Love & Turner, [6]). The speed of learning and acceptability of voice-activated menu interfaces are also related to the extent to which callers report visualising the phone service they are using, a task which becomes easier when spatial metaphor is used (Howell, Love & Turner, [7]). Logan, Sanders, Synder, Morris & Buckner, [8] showed that providing the elderly with appropriate processing strategies to aid recall on a memory task can be an effective way of partially offsetting age-related memory deficits. It follows that the provision of an effective metaphor for interacting with a phone based menu system may present a viable means of improving the accessibility of such telephone systems for older adults.

1.1 Aims

The aim of the current study was to examine whether a voice-activated telephone menu system designed around a spatial metaphor could be used to enhance the usability of telephone services for older adults, when factors such as cognitive ability and visual memory are taken into account.

Specifically, the study addresses:

1. Whether a spatial-metaphor based service would be easier and faster to use than a standard service that uses numbered menus.
2. Whether ability to visualize the structure of each service improves users' interactions with the service.
3. Whether individual differences in cognitive attributes such as working memory capacity and spatial ability, and previous experience of telecommunication systems can be used to predict success at using automated services.

2 Method

2.1 Design

The study used a 2 (trial) x 2 (service type) repeated measures design. A hierarchically structured phone menu system was used which consisted of five levels of service messages and prompts, with a maximum of three menu options to choose from at each level. Users completed two search tasks with each version of the service which required them to successfully retrieve up to four pieces of information before exiting the service.

The non-metaphor based version of the service was designed in the same style as commonly available automated phone services by pairing numbers with menu options (e.g. "Say '1' for Arts and Entertainment, say '2' for Eating out, etc.).

The office filing system service required the caller to select one of three filing cabinets and then to navigate through different drawers, partitions and folders to the information required. Commands within the system were designed to encourage spatial processing of the system structure (e.g. top/bottom, front/middle/rear, etc.). In addition, participants were provided with a cue card depicting three filing cabinets which provided a simple visual representation of the top level of the service, designed to facilitate visualization (Fig. 1).

Fig. 1. Cue card for metaphor-based service structure

2.2 Measures

The following measures were gathered from all participants in the study:

Task Performance Measures. Two objective measures of task performance were collected during each participant's interaction with the service; successful task completion and the time taken to complete each task. The total time taken to complete each task was logged in seconds from the beginning of the first 'level 1' dialogue to the end of the 'exit' dialogue.

Service Usability. A 20-item post-exposure usability questionnaire was used to evaluate subjective attitudes towards each implementation of the service, balanced for positively and negatively worded items. Items (e.g. "I sometimes felt lost when using the service") were responded to using a 7-point Likert scale, where higher scores indicated greater difficulty with using the service. In addition, participants were asked to indicate 'yes' or 'no' whether they had been able to visualize each service.

Technographic Questionnaire. This consisted of a series of 18 questions designed to assess each person's previous experience of using mobile telephones, automated telephone services and other relevant forms of technology. Items (e.g. "How often do you use a mobile telephone to access automated services such as banking or cinema information?") were responded to using a visual analogue scale with scores in the range 0 to 10, where higher scores indicated greater frequency of use.

Spatial Ability. The AH4 Group Test of General Intelligence (Heim, [9]) Part II was used to assess spatial ability. This is a two-section questionnaire consisting of 12 practice items and 65 test items in each section. Participants are presented with a target diagrammatic stimulus or series of stimuli and are required to perform visual analogies, subtractions or superimpositions in order to correctly identify an appropriate transformation of the original target. Participants are permitted a maximum of 10 minutes to complete Part II, resulting in a numerical indication of spatial ability between 0 and 65.

Working Memory Span Test. This test was used to assess participants working memory capacity derived from Baddeley [10]. Participants were serially presented with sentences which were either semantically correct or incorrect. Sentences were read aloud to participants. Each sentence consisted of 5 words (e.g. “The boy brushed his teeth”). After listening to each sentence the participant was asked to say ‘yes’ if the sentence made sense, or ‘no’ if the sentence did not make sense. After each block of 5 sentences, participants were asked to recall the last word of each of the 5 sentences in that block. Eight blocks of sentences were presented in total, yielding a maximum working memory score of 40 correctly recalled items.

2.3 Participants

Twenty-two individuals took part in the study consisting of 12 males and 10 females. The age of participants ranged from 50 to 87 years (M=61.0 years, SD=10.9) with participants being recruited from members of a local old people’s group and older working members of a local organization. The study was conducted in accordance with the British Psychological Society (BPS) Ethical Principles for Conducting Research with Human Participants as defined by the BPS Code of Ethics and Conduct [11] and was approved by the relevant Research Ethics Committee of our organization. All participants gave their consent to take part in this research and for their data to be used in this report.

2.4 Procedure

Each participant was tested individually within a 1-hour session. Firstly participants completed the technographic questionnaire, followed by the working memory span test and AH4 assessment of spatial ability. Each participant then completed one practice task and two recorded search tasks each with the standard service and the metaphor-based service respectively. A ‘Wizard of Oz’ methodology was used for the experiment which involved the experimenter simulating the functionality of a fully automated system to create the illusion that the user was interacting with a real voice-activated telephone service. Participants gave verbal rather than keyed responses and listened to the service prompts via a headset. Both services were based around the provision of information for a mock telephone city guide service (Howell et al, [6]). Each task required participants to find and retrieve specific information from within the guide (e.g. “Find the names of 2 weekend art exhibitions and then exit the service”). Task completion times were recorded using a stopwatch. On completion of the tasks within each service, participants completed the usability questionnaire.

3 Results

3.1 Task Performance Measures

Task Completion. Use of the automated telephone services presented difficulties for the current sample with only 13 out of the 22 participants able to successfully complete all tasks with standard service (59.1%) but significantly more (77.3%) were able to complete all tasks with metaphor-based service $\chi^2(1, n=22)=9.35, p=0.005)$.

Fig. 2 shows the mean task completion rate achieved by participants on the first and second trail with each service as a percentage of the total number of pieces of information that were to be retrieved.

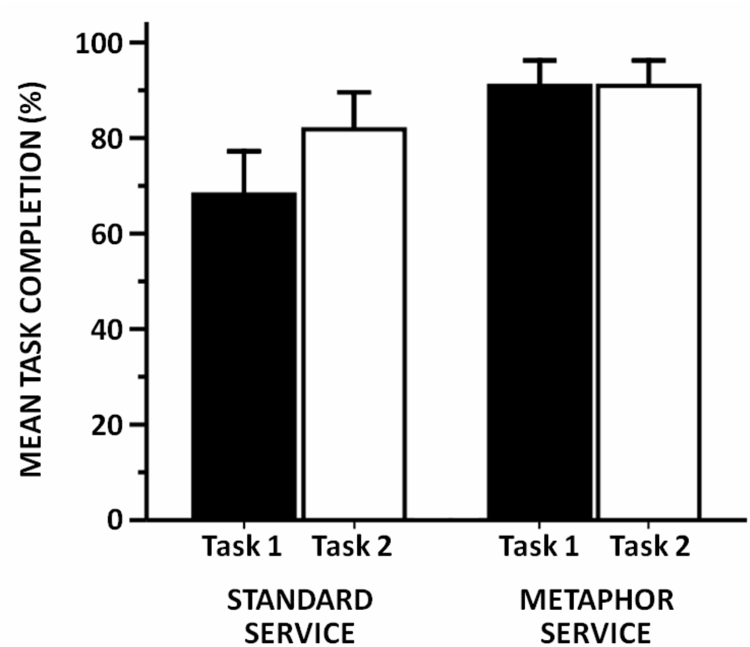


Fig. 2. Mean (+1 SE) successful task completion (percent correct information retrieved) for two calls made to a standard service and metaphor-based phone service by 22 older adults

A 2x2 (service x trial) repeated measures analysis of variance ANOVA revealed a main effect of metaphor ($F(1,21)=8.871$, $p=0.007$; $\eta^2=0.297$). Successful task completion was greatest for the metaphor service (90.9%, SD=25.1%) and lower for standard service (68.2%, SD=42.4%) on task 1 increasing to 81.8%, SD=36.3%) on task 2. Bonferroni Pairwise comparisons revealed a significant difference to occur between the metaphor and standard service on task 1, suggesting that information could be better retrieved from the metaphor service on first use. No main effect of task ($F(1,21)=2.278$, $p=0.110$ ns; $\eta^2=0.117$) and no interaction effect ($F(1,21)=2.278$, $p=0.110$ ns; $\eta^2=0.117$) were found suggesting that the difference in task completion rates between the two services remained broadly consistent across both tasks.

Task Duration. For the 13 participants who successfully completed all trials (Fig. 3), the completion times did not increase significantly from the first call to second call ($F(1,12)=1.85$, $p=0.198$ ns; $\eta^2=0.134$). There was also no overall difference ($F(1,12)=0.121$, $p=0.734$ ns; $\eta^2=0.10$) in the time taken to use standard service (M=2.7 minutes; SD=1.0) and metaphor service (M=2.8 minutes; SD=1.1). A significant interaction was found though between service type and trial ($F(1,12)=10.941$, $p=0.006$; $\eta^2=0.477$) which arose from the fact that less time was taken to use the standard service (M=2.6 minutes) than metaphor service (M=3.0 minutes) during performance of the second task (Bonferroni Pairwise comparison, $p=0.015$). This difference may have been an artifact of the specific information that participants were asked to seek which was varied across services to prevent rehearsal.

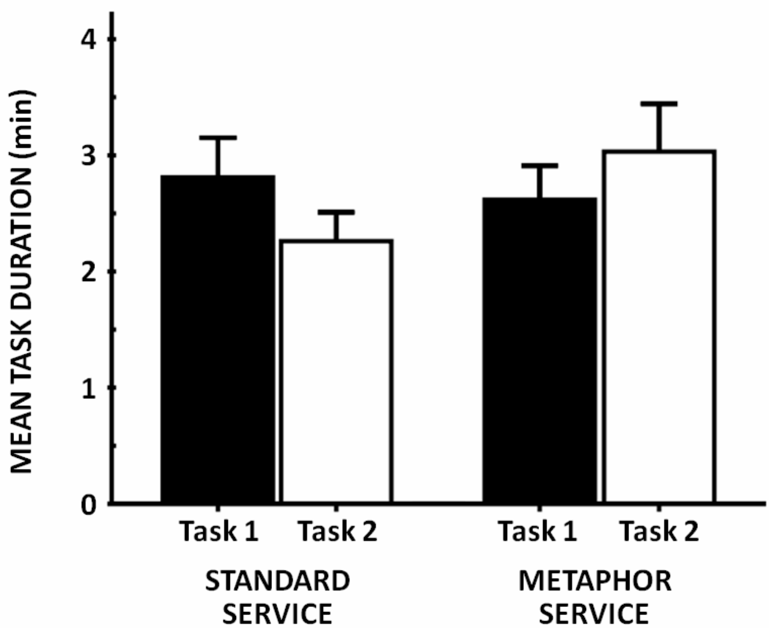


Fig. 3. Mean (+1 SE) task completion duration (minutes per call) for successful calls made to a standard service and metaphor-based phone service by 13 older adults

3.2 Service Usability and Visualization

The self-reported usability of the two telephone service indicated participants found both services difficult to use, with no overall difference ($t(21)=1.42$, $p=0.16$ ns; $d=0.43$) being found between the subjective evaluations of the standard service (M=3.91, SD=1.33) and the metaphor-based service (M=4.41, SD=1.02).

A greater number of participants reported being able to visualize the metaphor-based service (86.4%) than were able to visualize the standard service (68.2%). Visualization was also related to service usability. Participants who reported being able to visualize the standard service found this service significantly easier to use (M=4.32; SD=1.00) than those who could not visualize the service (M=3.04; SD=1.6) (t(20)=2.30, p=0.032; d=0.98). However, in the case of the metaphor-based service the difference in usability ratings between those who could (M=4.52, SD=1.00) and could not (M=3.75, SD=1.05) visualize the service was not found to be statistically significant (t(20)=1.23, p=0.23 ns; d=0.74).

3.3 Individual Difference Variables

Table 1 shows Pearson correlation coefficients between the two objective measures of task performance for both versions of the phone system with individual participant attributes including spatial ability and working memory capacity.

Table 1. Pearson (r) Correlation Coefficients (* p<0.05 ** p<0.01; 20df) between cognitive ability measures for 22 older adults and task performance measures when using the standard and metaphor-based services

Attribute	Standard Service		Metaphor-based Service	
	Task Completion	Task Duration	Task Completion	Task Duration
Age	-0.56**	-0.06	-0.67**	-0.01
Spatial Ability (AH4)	0.67**	-0.62**	0.61**	-0.33
Working Memory Span	0.54*	-0.59**	0.32	-0.20
Service Usability	0.47*	-0.50*	-0.29	0.06
Technographic experience	0.15	-0.05	-0.06	-0.17

These analyses suggest that successful task completion declined significantly with increased participant age for both the standard service (r(20)= -0.56, p<0.01) and the metaphor-based service (r(20)= -0.67 p<0.01). However, the time taken to complete tasks was unrelated to age for both services.

Participants with higher spatial ability scores were found to have better task completion rates for both the standard service (r(20)= 0.67, p<0.01) and metaphor-based service (r(20)= 0.61, p<0.01). Those with greater spatial ability also completed tasks more quickly with the standard service (r(20)= -0.62 p<0.01) although no significant relationship was found between spatial ability and speed of completion for the metaphor-based service (r(20)= -0.33, ns).

Working memory capacity was related to task performance for the standard service where by those with a greater working memory span for phonological information were able to complete tasks more quickly (r(20)= -0.59 p<0.01) and more accurately (r(20)= 0.54 p<0.05) with the standard service. However, working memory span was not related to the speed or accuracy of task performance when using the metaphor-based service.

Subjective ratings of service usability were also associated with quicker ($r(20)= -0.50$ $p<0.05$) and more successful task completion rates ($r(20)= 0.47$ $p<0.05$), but again this trend was only observed for the standard service and not for the metaphor-based service.

Participants' speed and accuracy of task performance was found to be unrelated to their previous experience of using automated telephone systems in the case of both the standard and metaphor-based phone services.

4 Discussion

Of the two telephone systems examined, the interface which used a spatial metaphor based around an office filing system was easier to visualize and resulted in more successful service use among older adults than a standard numbered menu system. The requirement to complete only 2 tasks with each menu system following only one practice trial was designed to mimic the actual context in which such systems would be used, where participants will often have little prior knowledge or exposure to menu options or structure. Under these conditions, when examining initial task performance with each system, successful task completion was found to be significantly better for the metaphor-based service. This result is consistent with previous research that suggests visual navigational aids can improve the usability of menu structures for older adults (Ziefle & Bay, 2006).

Contrary to expectations, the metaphor-based service did not result in faster task completion times than the standard service. Whilst the absolute time taken to complete tasks provides a valid indicator of the level of involvement with each service required by the user, it does not provide a true test of usability when comparing the two services directly since service prompts were of different lengths in each system. Since the metaphor-based service tended to have longer service prompts but did not produce significantly longer task completion times overall, this may provide a strong indication of the potential utility of this service. Numbered menu systems are well established in contemporary automated phone services. It could therefore be seen as encouraging that a novel system design with longer service prompts was not evaluated by older adults as being less preferred or indeed did not result in a significant increase in the time taken to use the service use overall.

One possible explanation for the superior user performance with the metaphor-based system is assumed to be the emphasis this service places on the use of spatial as opposed to verbal cues when interacting with the service. Those with greater spatial ability were found to perform more successfully on tasks with both services, which may indicate that the ability to imagine and manipulate the menu service structure within visual memory provides the key to better performance. Fewer participants reported being able to visualize the standard service structure and performance with the standard service only was found to be related to working memory span. This may indicate that in the absence of a reliable visual representation of a telephone system, service users are more reliant on phonological aspects of memory when searching serially presented verbal information, and that providing a suitable visual structure to represent phone systems to users may help alleviate the burden placed on the phonological loop component of working memory that is experienced by older adults when interacting with such systems.

Metaphor-based interfaces may therefore provide an important aid to automated service use for older adults and present a possibility for future system design. Further research might give consideration to the utility of alternative metaphors to support service structures beyond an office-based system that might benefit older adults or offer a broader, more universal design.

Acknowledgments. This research was supported by a grant from the British Psychological Society Undergraduate Research Assistantship Scheme.

References

1. Office for National Statistics. Population Estimates: UK population grows to 61.8 million (2009), `http://www.statistics.gov.uk/cci/nugget.asp?id=6` (retrieved February 2, 2010)
2. Renaud, K., Van Biljon, J.: Predicting Technology Acceptance and Adoption by the Elderly: A Qualitative Study. In: Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists, October 6-8 (2008), `http://researchspace.csir.co.za/dspace/bitstream/10204/2514/1/Kotze_2008.pdf` (Retrieved February 2, 2010)
3. Ziefle, M., Bay, S.: How to Overcome Disorientation in Mobile Phone Menus: A Comparison of Two Different Types of Navigation Aids. Human-Computer Interaction 21, 393–433 (2006)
4. Roring, R.W., Hines, F.G., Charness, N.: Age Differences in Identifying Words in Synthetic Speech. Human Factors: The Journal of the Human Factors and Ergonomics Society 49(7), 25–31 (2007)
5. Kausler, D.H., Salthouse, T.A., Saults, J.S.: Temporal memory over the adult lifespan. American Journal of Psychology 101, 207–215 (1988)
6. Howell, M., Love, S., Turner, M.: The impact of interface metaphor and context of use on the usability of a speech-based mobile city guide service. Behavior & Information Technology 24(1), 67–78 (2005)
7. Howell, M., Love, S., Turner, M.: Visualization improves the usability of voice-operated mobile phone services. Human-Computer Studies 64, 754–769 (2006)
8. Logan, J.M., Sanders, A.L., Snyder, A.Z., Morris, J.C., Buckner, R.L.: Under-Recruitment and Non-Selective Recruitment. Neuron 33(5), 827–840 (2002)
9. Heim, A.W.: AH4 Group Test of Intelligence Manual. National Foundation for Educational Research, Slough (1967)
10. Baddeley, A.: Working Memory. University Press, Oxford (1986)
11. The British Psychological Society, Code of Ethics and Conduct: Guidance published by the Ethics Committee of the British Psychological Society. BPS, Leicester (2009)

Oracle Mobile User Assistance Testing

Erika Webb[1], Ultan O'Broin[2], Ray Matsil[3], Laurie Pattison[3], and Mindi Cummins[3]

[1] Oracle Applications User Experience, 500 Eldorado Blvd. Broomfield CO 80027 USA
[2] Oracle Applications User Experience, Eastpoint Business Park Alfie Byrne Road Block P5 Dublin, Dublin 3 Ireland
[3] Oracle Applications User Experience, 300 Oracle Parkway, Redwood City , CA 94065 USA
{erika.webb,ultan.obroin,ray.matsil,laurie.pattison, mindi.cummins}@oracle.com

Abstract. In order to create a set of Oracle Mobile Applications patterns and guidelines that would reflect the needs of customers both in the US and other global areas, we conducted a set of user feedback sessions on messages with mobile application users in both the UK and the US. Participants were asked to consider scenarios that covered how they expected notifications to work, collaboration scenarios, how messages might display, specific wording alternatives for messages, and expectations for confirmation messages. The information gained from these interviews and sessions was used to validate assumptions, previous research findings, and anecdotal evidence, as well as to explore wants and needs for future releases of mobile applications for enterprise users. This information was then used to create a new set of patterns for mobile messaging, as well as to create guidelines for content of messages for our development teams.

Keywords: Mobile, Messaging, User Assistance, patterns, guidelines.

1 Introduction

In early 2009, the Oracle Mobile User Experience team developed and tested a set of mobile design patterns for Oracle mobile applications. Because there has been a significant surge in smart phone usage and in smart phone technologies, the Oracle Applications User Assistance decided to revisit some of the previous findings, as well as test new areas around message language and management. The Oracle User Assistance group was also interested in seeing whether there was a difference between user preferences in the UK and the US. A set of 11 semistructured interviews was conducted in the Oracle Usability Lab in Thames Valley Park (TVP), Reading, UK. A second set of 10 interviews was conducted during the Oracle Usability Lab sessions at Oracle OpenWorld (OOW) 2010. Participants were asked to consider scenarios that reviewed the following:

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 687–695, 2011.
© Springer-Verlag Berlin Heidelberg 2011

- Notifications
- Collaboration scenarios
- Message display options
- Specific wording alternatives
- Confirmations

The information gained from these interviews and sessions was used to validate assumptions, previous research findings, and anecdotal evidence, as well as to explore wants and needs for future releases of mobile applications for enterprise users. This information was then used to create a new set of patterns for mobile messaging, as well as to create guidelines for content of messages for our development teams.

All of the sessions were semistructured interviews in which participants were shown mock-up designs and asked to discuss their feelings and preferences about notifications, collaboration, message display options, wording, and confirmations. The screens were low-fidelity designs. Participants were asked to focus not on how their phone worked, but how they wanted an application on a phone to work.

2 Participants

A total of 21 participants were recruited for participation in the studies in the UK and the US. All participants were smart phone users who used some type of application on their phones for work (in addition to email and internet browsing applications). Table 1 shows participants from the UK, including their job titles and mobile devices.

Table 1. Participants from the UK

Role	Devices and Operating Systems
Oracle HR Support and Delivery Manager	iPhone 3gs
Master Principal Sales Consultant	iPhone 3gs
Principal	Palm Treo Pro (WinMobile 6.1)
AMS Manager	Palm Treo Pro (WinMobile 6.1)
Automation Director, Customer Intelligence, Global Support	HTC HD2 (WinMobile 6.5)
Group Leader Networks and Communications	HTC HD2 (WinMobile 6.5) Blackberry 8707
Group Vice President	Blackberry
VP, Corporate Communications EMEA	Nokia N97 (Symbian 9.4)
Architect	HTC Desire (Android 2.1)
Business Development Manager	Google Nexus One (Android OS 2.0)
Director, Product Management	HTC Desire (Android 2.1)

Table 2 shows the job titles and devices of participants in the US study.

Table 2. Study participants from the US and their mobile devices

Role	Devices and Operating Systems
HCM Systems Manager	iPhone 3gs
Sr. IT Business Analyst	iPhone 3g
Director, Information Systems	iPhone 3g
Sr. Oracle Apps DBA	iPhone, iPad
Managing Partner	iPhone 4
CTO & Founder	iPhone 4/Motorola Droid (Android 2.0)
Chief Information Officer	iPhone, iPad, Droid (Android 2.0)
Executive Director, Product Development	iPhone, Blackberry Bold
Senior Consultant	Blackberry + iPod Touch
CEO	Blackberry Bold

3 Tasks and Findings

3.1 Notifications

Notifications inform users that some event has occurred in their applications or at the level of the phone system. For example, an alert about a missed call or new voice mail is a type of notification. In this study, we were interested in better understanding how users of Oracle products wanted to receive notifications of activity from an Oracle mobile application.

Participants in the study were shown three design alternatives for notifications about events in an Oracle mobile application. In the first scenario, participants were shown a set of universal notifications, where anything that happened on the phone appeared in a single location (for example, missed calls). In the second scenario, all Oracle mobile application notifications appeared grouped together by application, but separated from other activity on the phone (for example, new voicemail). In the third scenario, Oracle mobile application notifications appeared only in the specific application. In this scenario, the user would have to open the specific application in order to see the notification.

In this set of scenarios, we found a difference in user preference by region. Participants in the UK primarily preferred the option to have all notifications grouped in a single location. On further discussion, the UK participants indicated that they wanted a single Oracle notification line in the global list that would drill down into just the Oracle notifications (essentially a combination of A and B).

Participants from the US at OOW were split between preferring to have all notifications for anything on their phones grouped together and preferring to have all Oracle applications grouped together. Participants who preferred the second option seemed to be generally concerned about the number of notifications that they receive on their phones. These participants were worried about losing their Oracle notifications in the long list of notifications that appeared. These participants felt that they would rather have the Oracle applications notifications separated from their other notifications.

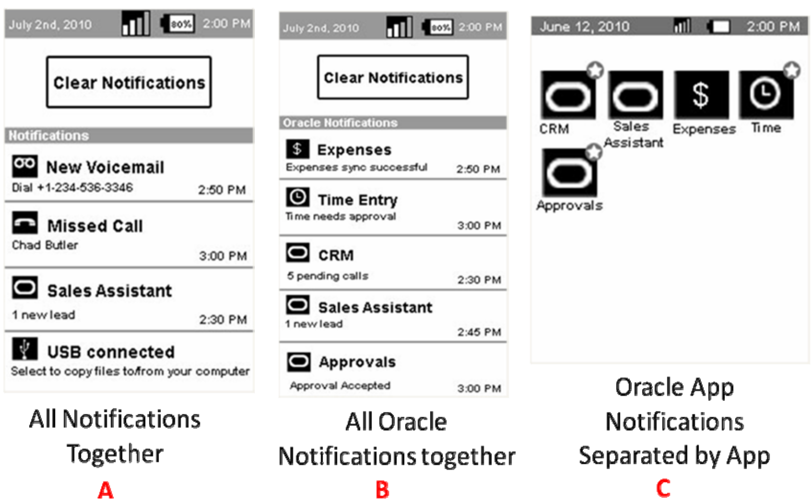


Fig. 1. The three notification scenarios participants evaluated

Our recommendation to our development teams was to surface notifications or alerts from Oracle mobile applications at the level of the system applications. However, if the user has multiple Oracle mobile applications, development should consider a single notification point in the global notifications that the user can drill into in order to see individual application notifications.

3.2 Collaboration

In the case of collaboration alternatives, we were interested in different kinds of collaboration scenarios. In the first scenario, participants had an option to share information with a coworker. In the second scenario, participants were shown an option to share a problem with user support.

Collaborating with a Coworker. In the scenario involving sharing with a coworker, we showed the example of a Customer Relationship Management (CRM) application in which a sales person was able to share information about a particular opportunity with a coworker (see Figure 2). In this particular scenario, it was explained that the information was more than sharing just a screen from the app, but in fact the act of sharing opened all of the sales opportunity information in the back end CRM system to the coworker.

Participants in both the UK and the US felt that collaborating with a coworker could be useful. However, some participants were unsure whether they would need this functionality for the type of work that they did. Participants felt that the sharing option would be especially useful for CRM, where users have a need to involve coworkers in their work. In addition, participants in the study in both the US and the UK commented that it would be useful to collaborate at a more granular level, such as assigning tasks, adding multiple users, and tagging items or areas for particular users to review.

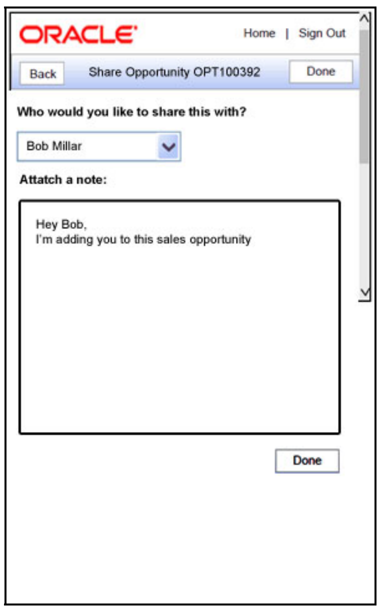


Fig. 2. A collaboration scenario in which a user would share information (in this case, a sales opportunity) with a coworker from within the mobile application

Collaborating with Support. In the sharing-with-support scenario, when an application error occurred, participants were given the opportunity to share that error with support. Participants in both the US and the UK first wanted a description of the problem to see if they could fix the problem themselves. UK participants preferred setting up a time for support to contact them rather than filing a ticket.

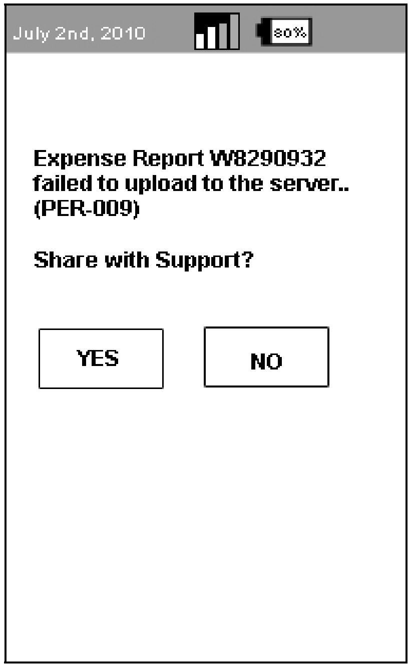


Fig. 3. Collaboration with support

3.3 Message Display Options

Participants were next asked to consider the ways that an error message might appear. In the first case, although there was an inline message with the error information, there was no field-level indication of the error. In option B, there were three different indicators that an error had occurred: a link to the error field, an error icon at the field, and a description of what the user would need to do to correct the error. Participants were asked to consider the usefulness of each of the three error indicators: the error icon, the error text, and the link to the location on the screen of each error.

Participants in both the UK and the US felt the red X error icons and instructional text were useful indications of what the user would need to do to correct the error. However, in both countries, the users in this study did not feel that the link was necessary. Participants indicated that they thought that mobile application screens would not be as long as screens on a desktop, laptop, or web application, and therefore, the link that would take them directly to the error field was not as critical.

Our recommendations for errors were to indicate the number of errors with an error icon and informational text in proximity to the field that needed to be corrected.

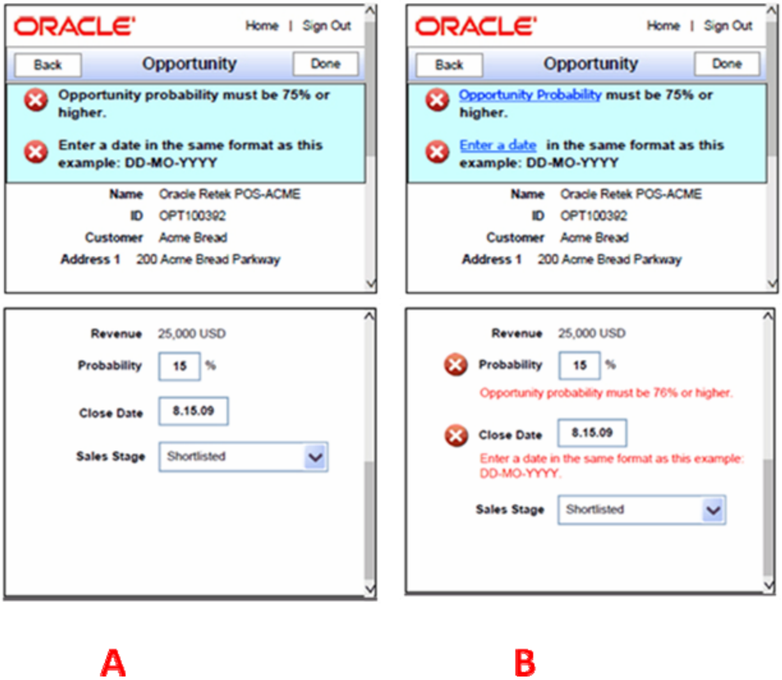


Fig. 4. The error message display option scenarios

3.4 Wording

Participants were asked to consider five different options for the wording of a confirmation message. The options included more active and passive phrases, different details, and different levels of formality (see Figure 5.)

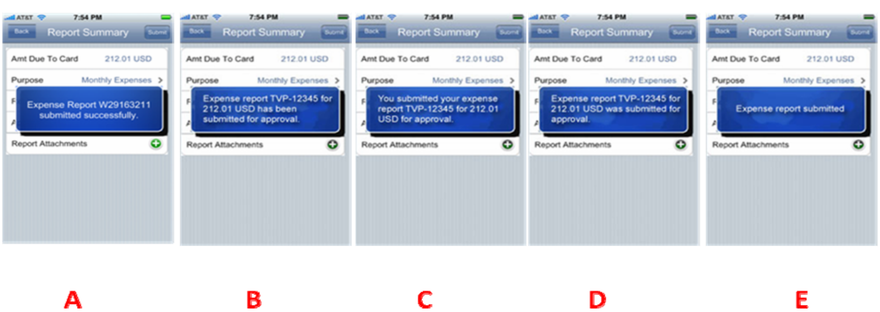


Fig. 5. The five wording alternatives presented to participants

None of the participants liked the use of “you” and “your” (“You submitted your expense report TVP-12345 for 212.01 USD for approval”). Several commented that this wording seemed too “cute” or informal for an enterprise mobile application.

UK. Participants in the UK were split on the need for the word “successfully” as part of the message. Some felt that it did add a sense of confirmation that the action completed, while others felt that the expense report number was a more important confirmation that the report had been sent to the server.

Overall, UK participants preferred options A and B, about equally. They felt that the most essential information was the report number and the amount.

US. Half of the participants in the US preferred option E (“Expense report submitted”), while the rest of the participants were split across the other options. The majority (7/10) felt that the word “successful” was important to include in the message. These participants were split on the need for the report number to be included in the message.

According to one participant, “*Successfully* means it was accomplished; that’s important.”

3.5 Confirmations

Users were asked to consider two different styles of action confirmation. In option A, the confirmation appeared as a pop-up message, while in option B, the confirmation appeared as an inline message.

Both the UK and US participants preferred inline confirmations. The participants said that it was useful to see the data and the message at the same time and preferred that the message remain persistent because it matched the way that they use their phones.

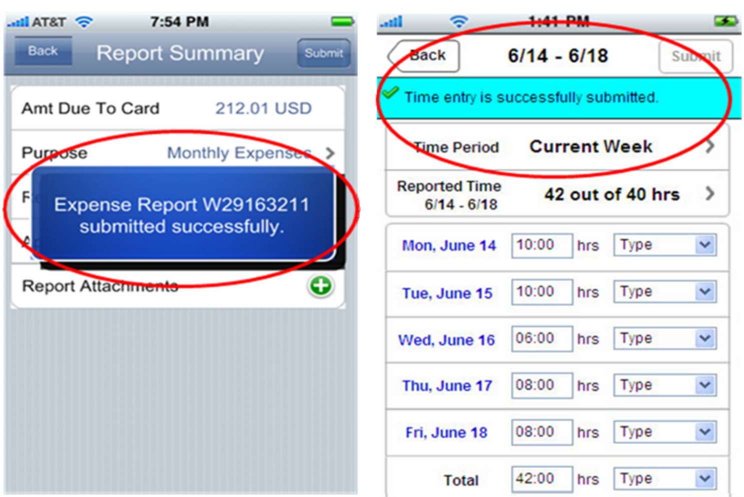


Fig. 6. Pop-up versus inline confirmation message options

4 Conclusion

Overall, this study found few differences between participants in the UK and the US in their preferences in mobile applications. The results were used to update and extend mobile messaging patterns and guidelines for Oracle mobile applications. In addition, further research is under consideration for other areas of Europe and Asia Pacific to ensure that these patterns and guidelines represent the needs and wants of a more global audience.

Proposal for Indices to Assess Attractiveness on Initial Use of Mobile Phones

Misako Yamagishi[1], Hideo Jingu[1], Keiko Kasamatsu[2], Hiroaki Kiso[3], and Shinichi Fukuzumi[3]

[1] Kanazawa Institute of Technology,
3-1, Yatsukaho, Hakusan-city, Ishikawa, 924-0838, Japan
[2] Tokyo Metropolitan University,
6-6, Asahigaoka, Hino-city, Tokyo, 191-0065, Japan
[3] Information and Media Processing Laboratories, NEC Corporation
1753, Shimonumabe, Nakaharaku, Kawasaki-city, Kanagawa, 211-8666, Japan
yamagishi-m@neptune.kanazawa-it.ac.jp,
jinguh@neptune.kanazawa-it.ac.jp,
kasamatu@sd.tmu.ac.jp, h-kiso@ah.jp.nec.com,
s-fukuzumi@aj.jp.nec.com

Abstract. It is necessary to determine the attractiveness level of using mobile phone to ensure user satisfaction. This study measured physiological indices of attractiveness during participants' initial use of a mobile phone. As physiological indices which concern autonomic nervous system activity, nasal skin temperature, pupil diameter, electroencephalography, blinking, and electrocardiography are listed. These indices were selected because preceding studies have suggested "safety and relaxing" and "feelings induced by desire and interest" as factors related to the attractiveness of mobile phones. The results obtained in these experiments that a device's attractiveness to users can be evaluated using physiological indices. Thus, the present study showed basic perspectives related to attractiveness from the point of physiological response.

Keywords: attractiveness, physiological measurement, mobile phone.

1 Introduction

Information devices are becoming increasingly advanced with progress in science and technology, making our lives more convenient. Although such advanced devices are attractive, it is difficult to satisfy users. It is necessary to determine the level of attractiveness of mobile phones to users to ensure user satisfaction. As a fundamental study, the present experiment aims to determine how to measure the level of attraction users feel toward mobile phones during their initial use by using objective indices such as physiological indices.

Studies regarding attractiveness-related *Kansei* elements have found that attractiveness may be associated with a sense of safety and relaxation induced by satisfaction or pleasantness and desire, interest, and emotion. On the basis of these findings, this study suggests that attractiveness can be assessed by indices that reflect autonomic nervous system activity [1] [2].

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 696–705, 2011.
© Springer-Verlag Berlin Heidelberg 2011

Hence, the present study considered that the attractiveness level with respect to induced interest and emotion as well as a safety and relaxation can be measured by adopting nasal skin temperature, pupil diameter, electroencephalography (EEG), blinking [3], and electrocardiography (ECG) as indices related to autonomic nervous system [4]. In particular, the study emphasized ensuring usuality in the measurement situation and a non-invasive, non-contact method. The experiment examined whether physiological indices can be considered to be objective indices to measure attractiveness by adopting the method of minimized body contact or constraint.

2 Experimental Methods

2.1 Measurement of Nasal Skin Temperature and Pupil Diameter

Eleven undergraduate and graduate participants (six men and five women, mean age = 22.1 years, SD = 1.30) participated in the experiment to measure nasal skin temperature and pupil diameter.

Nasal skin temperatures were obtained using thermography (infra-eye 3000, NIHON KOHDEN Corp.). Thermal images of facial surface were recorded every minute by an infrared ray detector. From the recorded images, area ranging from under the eyes to the upper lip (vertically) and from the right to left edge of the face (horizontally) was specified. Maximum and minimum temperatures in the area were determined, and the temperature difference between them was calculated.

The pupil response was obtained using an Eyemark recorder (nac Image Technology Inc.) having a sampling rate of 60 Hz. Pupil diameters were analyzed using "dfactory". Pupil diameters obtained here were transformed into a standardized *Z* score as an analytical indicator.

2.2 Measurement of EEG

Ten undergraduate and graduate participants (eight men and two women, mean age = 21.7 years, SD = 0.67) participated in the EEG experiment.

Exploring electrodes were placed on Fz, Cz, Pz, and Oz recording sites in accordance with the international 10/20 system [5]. Reference and ground electrodes were placed at C3 and C4 and both ear lobes, respectively. EEG was amplified using a digital EEG (Neurofax 1100, NIHON KOHDEN Corp.). The computer sampled at a rate of 1000 Hz. Fig. 1 illustrates the electrode locations in this experiment, as observed from above the head.

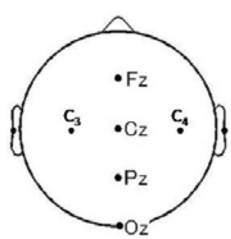


Fig. 1. Location of electrodes for EEG in this experiment

2.3 Measurement of Blinks and ECG

Ten undergraduate and graduate participants (10 men, mean age = 21.2 years, SD = 0.63) participated in the experiment to measure blinks and ECG.

Blinks were obtained by taking an eye image from a video camera per 1/30 frame. By observing images of each participant, two experimenters counted the number of blinks during each experimental operation of a mobile phone, and the average was calculated using the number of blinks obtained in both experimenters.

ECG was obtained using an RF-ECG wireless sensor (Medical Electronic Science Institute Co., Ltd.). The two electrodes were placed on the participant's left chest. The sampling rate was 204 Hz. RR intervals from an ECG signal were computed, and then spectral analysis was performed using FFT to compute the Low Frequency (LF) and High Frequency (HF) components. The LF/HF ratio was calculated as a measure of autonomic nervous system activity.

2.4 Experimental Procedure

First, participants gave their informed consent to take part in the study and the experimental procedure was explained to them in detail. They were given a scenario in which they were asked to evaluate a new phone as a possible replacement to their current mobile phone. On the basis of the preceding experimental results, three mobile phones were chosen: Models A, B, and C. The features of each are described in detail below. Model A has a touch screen and can be controlled by an onscreen software keyboard, except for a key to return to the initial screen; model B can be controlled by hardware keys (HW key) similar to some existing models and has a specific screen design; and model C has not only a touch screen but also a pointing device and some HW keys (such as an on/off hook and menu etc), unlike model A.

The participants completed three common tasks using each of these models: 1) composing a new mail, 2) setting the alarm clock for 5:00 a.m., and 3) capturing three pictures. They were handed one of the three models and asked to perform the three tasks in order. After completing all the tasks, they were handed the remaining models. The same procedure was repeated for each of the three models. Finally, the participants were asked to evaluate each model for the degree of "Attractiveness," "Desire to own," "Desire to use," and "Desire to buy" on a scale of 0 to 10 and to rank the three models. The models were randomly handed to the participants. Fig.2 shows an example of the experimental procedure.

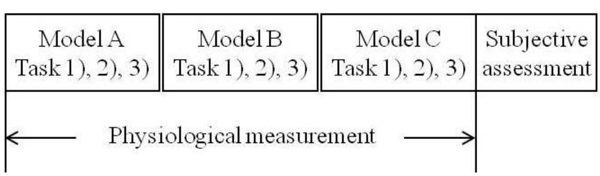


Fig. 2. Example of experimental procedure

3 Results

3.1 Nasal Skin Temperature and Pupil Diameter

For analyzing nasal skin temperature and pupil diameter, 10 participants, excluding a man whose data was incomplete, were assessed. Models A, B, and C were compared in terms of participants' physiological response while performing the task.

A one-way ANOVA was conducted for analyzing nasal skin temperature. Measurements of nasal skin temperature showed no significant differences in facial temperature between the models ($F(2,27) = 0.81$, *ns*), although facial temperature for model A is slightly higher than that for models B and C.

To examine differences between the models, a one-way ANOVA of pupil diameter was performed using standardized *Z* scores of pupil diameters. The analysis showed a significant difference among the models ($F(2,18) = 3.78$, $p < .05$). The multiple comparison by LSD method indicated that participants' pupils tended to be larger when they used model A or B than when they used model C. Fig. 3 shows the mean *Z* score of pupil diameter for each model.

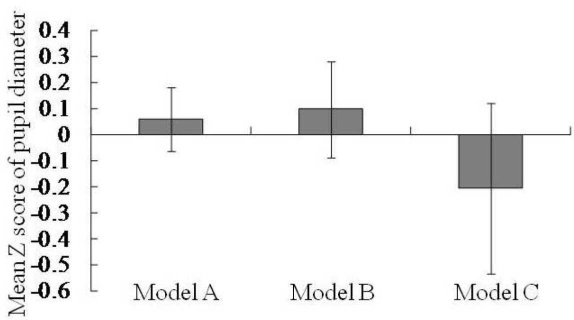


Fig. 3. Comparison of mean *Z* scores of pupil diameters of models

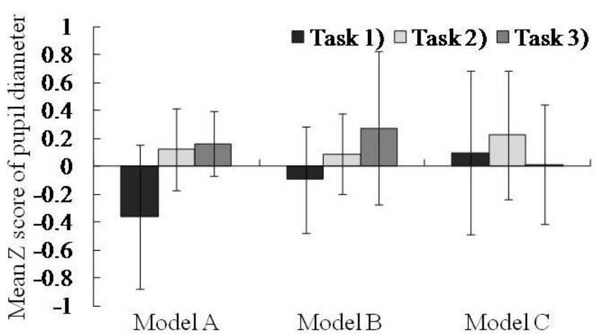


Fig. 4. Comparison of mean *Z* scores of pupil diameters

To compare the differences between tasks, the study used a two-way ANOVA for model and task. Results showed that the main effect of task is significant ($F(2,18) = 4.33$, $p < .05$). The results of multiple comparison by LSD method indicated that the pupil diameter in task 1) is smaller than in tasks 2) and 3). Fig. 4 shows the mean Z score of pupil diameter for each model.

3.2 EEG

The EEG analysis assessed nine participants, excluding a woman, because of noise in the data. The analysis examined differences between models in the amount of alpha wave (band of 8–13 Hz) associated with a state of relaxation and arousal. One-way ANOVA was performed using the amount of alpha wave for each of Fz, Cz, Pz and Oz. Results showed that there were no significant difference among the models for any the locations (Fz: $F(2,16) = 1.30$, *ns*; Cz: $F(2,16) = 1.11$, *ns*; Pz: $F(2,16) = 1.27$, *ns*; Oz: $F(2, 16) = 1.27$, *ns*).

Subsequently, two-way ANOVA for model and task was performed using the amount of alpha wave for Fz, Cz, Pz and Oz, separately. Measurements at Fz showed no difference in the amount of alpha wave among the models ($F(3,32) = 1.16$, *ns*).

For Cz, the main effect of task and the interaction of model and task are shown as follows: ($F(2,16) = 12.69$, $p<.01$; $F(4,32) = 3.37$, $p<.05$, respectively). The results of the test of simple main effect showed that for task 1), the amount of alpha wave of model A is smaller than that in of model C; for model A, the amount in task 1) is smaller than those in tasks 2) and 3); for models B and C, the amounts in tasks 1) and 3) are smaller than that in task 2).

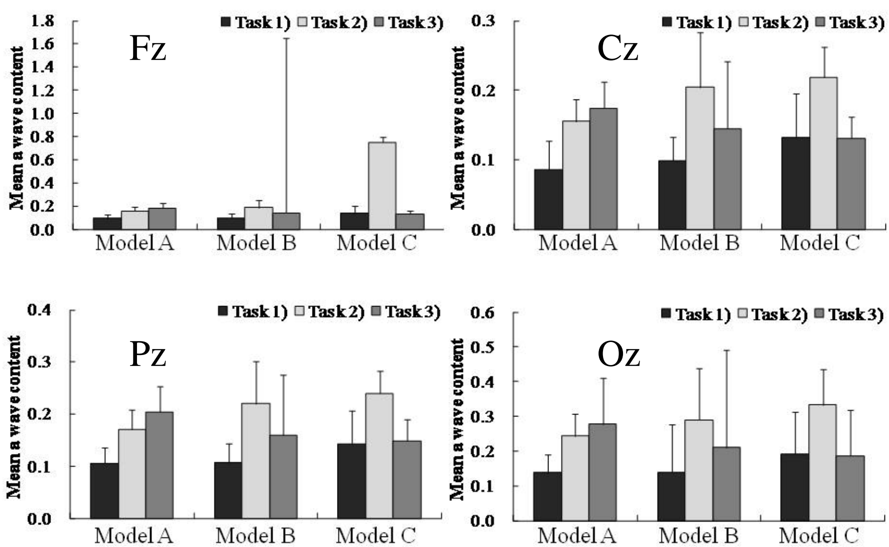


Fig. 5. Results of mean alpha wave amount for Fz, Cz, Pz, and Oz

Table 1. ANOVA tables for Fz, Cz, Pz, and Oz (**: p<.01, *: p<.05, *ns*: non-significant)

Fz

	SS	*df*	*MS*	*F*	*p*
Model	0.69	2	0.35	1.25	*ns*
Task	1.01	2	0.51	1.82	*ns*
Interaction	1.35	4	0.34	1.16	*ns*
Error	9.32	32	0.29		
Total	23.55	80			

Cz

	SS	*df*	*MS*	*F*	*p*
Model	0.00	2	0.00	0.53	*ns*
Task	0.11	2	0.05	12.69	**
Interaction	0.04	4	0.01	3.37	*
Error	0.09	32	0.00		
Total	0.40	80			

Pz

	SS	*df*	*MS*	*F*	*p*
Model	0.03	2	0.01	0.55	*ns*
Task	0.53	2	0.26	15.49	**
Interaction	0.20	4	0.05	3.52	*
Error	0.46	32	0.01		
Total	2.56	80			

Oz

	SS	*df*	*MS*	*F*	*p*
Model	0.01	2	0.00	0.29	*ns*
Task	0.24	2	0.12	8.42	**
Interaction	0.05	4	0.02	3.05	*
Error	0.22	32	0.01		
Total	2.01	80			

For Pz, analytical results indicated the main effect of task and the interaction of model and task: ($F(2,16) = 13.90$, $p < .01$; $F(4,32) = 3.24$, $p < .05$, respectively). The test of simple main effect revealed that for task1), the amount of alpha wave of model A is smaller than that of model C; for model A, the amount in at task 1) is smaller than that in tasks 2) and 3); for models B and C, the amount in tasks 1) and 3) are smaller than that in task 2).

For Oz, the main effect of task and the interaction of model and task are as follows: ($F(2,16) = 8.42$, $p < .01$; $F(4,32) = 3.05$, $p < .05$, respectively). The test of simple main effect indicated that for model A, the amount of alpha wave in task 1) is smaller than those in tasks 2) and 3); for models B and C, the amount in tasks 1) and 3) are smaller than that in task 2). Fig. 5 shows the results of amount of alpha wave for Fz, Cz, Pz, and Oz. The ANOVA results are summarized in Table 1.

3.3 Blinking and ECG

To compare the models, the study conducted an analysis of blinking and ECG by using the blink rate (blinks/second) during the performed task and the LF/HF ratio computed from the ECG data.

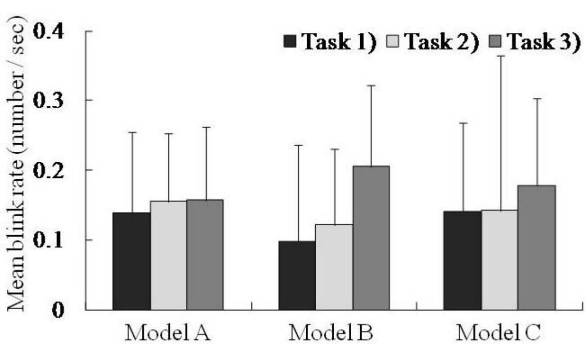


Fig. 6. Comparison of blink rate in models and tasks

A one-way ANOVA of blink rate was conducted to examine differences between models. The result showed no significant differences ($F(2,18) = 0.07$, *ns*). To examine the effects of task, two-way ANOVA for model and task was performed. According to the results of multiple comparison, the main effect of task indicated that the blink rates in tasks 1) and 2) were lower than that in task 3). The results for blink rates are shown in Fig. 6.

The LF/HF ratio was examined next using the Friedman test, which showed that there were no significant differences ($S(2) = 2.89$, *ns*). Two-way ANOVA for model and task showed the main effect of model and task ($F(2,18) = 3.00$, $p < .10$; $F(2,18) = 2.71$, $p < .10$, respectively). Multiple comparison of the effect of model indicated that the LF/HF ratio of model A was higher than that of model B. The comparison of task showed no differences among the models. The results for the LF/HF ratio are shown in Fig. 7.

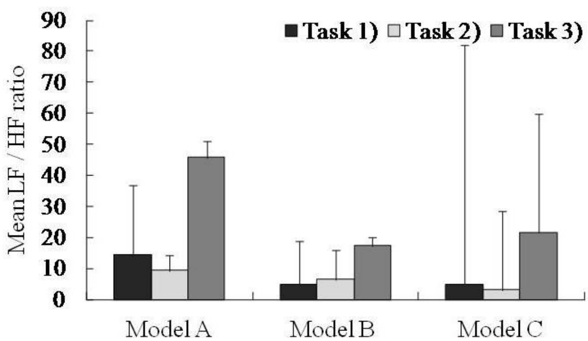


Fig. 7. Comparison of LF/HF ratio in models and tasks

3.4 Subjective Assessment and Model Ranking

The differences among models were examined in terms of the degree of "Attractiveness," "Desire to own," "Desire to use," and "Desire to buy" using separate one-way ANOVAs. It was confirmed that for all subjective assessments, model A was evaluated as significantly better than models B and C (Attractiveness: $F(2,60) = 19.74$, $p < .01$; Desire to own: $F(2,60) = 17.73$, $p < .01$; Desire to use: $F(2,60) = 14.36$, $p < .01$; Desire to buy: $F(2,60) = 19.55$, $p < .01$). The mean scores for the four subjective assessments are shown in Fig.8.

The ranks of the three models were determined using the Friedman test. The test revealed significant differences among the models ($S(2) = 24.47$, $p < .01$). The results of multiple comparison indicated that model A was ranked higher than models B and C.

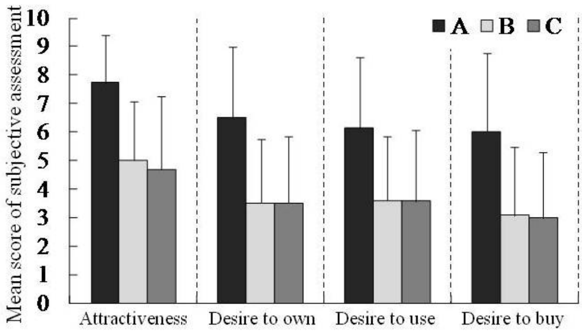


Fig. 8. Result of four subjective assessments of each model

3.5 Relationship between Physiological Indices and Subjective Assessments

To examine the relationship between physiological indices and subjective assessments, correlation coefficients were calculated. The results revealed significant correlations between pupil diameter and "Attractiveness" for models A and B. in addition, pupil diameter and "Desire to use" and "Desire to buy" tended to be correlated for model B, and blink rate and "Attractiveness" tended to be correlated for model C. Table 2 lists the correlation coefficients for physiological indices and subjective assessments.

Table 2. Relationship between physiological indices and subjective assessements

	Model A			Model B			Model C		
	Nasal skin temperature	Pupil diameter	Blink rate	Nasal skin temperature	Pupil diameter	Blink rate	Nasal skin temperature	Pupil diameter	Blink rate
Attrac-tiveness	−0.508	**0.708**	0.539	0.469	**0.675**	0.273	−0.268	0.013	−0.615
Desire to own	−0.371	0.169	0.425	0.375	0.511	0.242	−0.088	−0.16	−0.392
Desire to use	−0.182	−0.009	0.486	0.296	0.565	0.405	−0.206	−0.166	−0.092
Desire to buy	−0.393	−0.038	0.412	0.358	0.571	0.383	−0.373	−0.199	0.113

■ $p < .01$, ■ $p < .05$, ■ $p < .10$

4 Discussion

The purpose of this study is to find an objective index that expresses the attractiveness of a mobile phone during its initial use. To achive this aim, the present study carried out experiments that measured nasal skin temperature, pupil diameter, EEG, blink and ECG by considering the factors of "safety and relaxing" and "feelings induced by desire and interest" associated with attractiveness.

To examine differences in user reactions during use of mobile phones, physiological response was analyzed as participants completed three tasks: 1) composing a new mail, 2) setting the alarm clock, and 3) capturing pictures. The analysis of pupil diameter indicated a significant variation in size among participants‘ pupils as thet used the different models. The results show little evidence of substantial differences in nasal skin temperature, EEG, and blinking and ECG among the models. The changes in pupil diameter appear in response to emotional changes. Some studies have reported that because the size of the pupil is regulated by the autonomic nervous system, emotional stimuli such as interesting pictures evoke larger pupil dilation than neutral stimuli[6]. The present results showed larger pupil diameter in models A and B compared to model C. These differences may occur as a result of evoked interest and induced activity of the sympathetic nervous system for models A and B.

It was confirmed by subjective assessment that model A had higher scores in attractiveness and the other desires. In addition to the relatively larger pupil diameter, model A showed a higher LF/HF ratio. The results of the experiments suggest that physiological response and subjective assessment partially corresponded to the degree of attractiveness or desire evoked by these physiological responses, because the LF/HF ratio was associated with sympathetic nervous system activity [7] such as pupil diameter. This finding may indicate that attraction can be measured using pupil diameter and ECG.

Further studies on the combined effect of these indices are needed to establish the validity of physiological indices as measurements of attractiveness. If both indices reflect attractiveness, a similar tendency among models should be appear for pupil diameter and ECG. Therefore, it is important to measure pupil response and ECG simultaneously and to examine their association with attractiveness (subjective assessment).

Future study should focus on analysis of each task in detail. This suggestion is based on data that indicates user responses for all models tended to differ in some physiological indices depending on the task performance. This finding may indicate that the features of task and the process of operation related to task achievement might influence the measurement of attractiveness. More effort will be required to investigate task and function and to analyze the system and process of operation to expand the physiological index that expresses a products‘ attractiveness during its use.

5 Conclusion

As a fundamental study to obtain the possibility of objective indices related to a mobile phones‘ attractiveness during its initial use, the present experiment was designed using nasal skin temperature, pupil diameter, EEG, blinking and ECG. The results have shown the perceptions of different factors and physiological responses related to attractiveness.

The present study addressed mobile phone use. There is a need for further research that investigates various products to identify the relationship between a products‘ attractiveness during its use and users‘ physiological responses as objective indices. Regarding applications for engineering and high quality service, it would be valuable to develop technology for evaluating and measuring product attractiveness using a non-invasive, non-contact method and to explore practical applications of such technology.

References

1. Kiso, H., Kashiwase, N., Kasamatsu, K., Yamagishi, M., Jingu, H., Fukuzumi, S.: Relations between *Kansei* Elements of Attraction on Initial Use of Mobile Phone. The Japanese Journal of Ergonomics 46, 360–361 (2010) (Japanese)
2. Kashiwase, N., Kasamatsu, K., Kiso, H., Yamagishi, M., Jingu, H., Fukuzumi, S.: The Consideration for *Kansei* Factors of Attraction on Initial Use of Mobile Phone. The Japanese Journal of Ergonomics 46, 358–359 (2010) (Japanese)
3. Barreto, A., Zhai, J., Rishe, N., Gao, Y.: Significance of Pupil Diameter Measurements for the Assessment of Affective State in Computer Users. In: Advances and Innovations in Systems, Computing Sciences and Software Engineering, pp. 59–64 (2007)
4. Tada, H.: Eyeblink Rates as a Function of the Interest Value of Video Stimuli. Tohoku Psychologica Folia 45, 107–113 (1986)
5. Japanese Society of Clinical Neurophysiology: Guidelines for ClinicalEvoked Potential Studies. Japanese Journal of electroencephalography and electromyography 25(3), 1–16 (1997) (Japanese)
6. Partala, T., Surakka, V.: Pupil Size Variation as an Indication of Affective Processing. International Journal of Human-Computer Studies 59, 185–198 (2003)
7. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology: Heart rate variability; Standards of measurement, physiological interpretation, and clinical use. European Heart Journal 17 (3), 354-381 (1996)

Smart User Assistance Based on Dynamic Model Composition*

René Zilz and Peter Forbrig

University of Rostock, Department of Computer Science
Albert-Einstein-Str. 21, 18059 Rostock, Germany
{rene.zilz,peter.forbrig}@uni-rostock.de

Abstract. Ubiquitous computing environments are complex systems since they support a variety of different input modalities such as voice, touch and keyboard/mouse input. Especially inexperienced users (e.g., elderly or novel) might be confused with the abundance of interaction opportunities. This paper presents an approach to specify smart user assistance for mobile devices. The mobile application shows performable tasks, explains modalities and commands, and provides interactive exercises in order to get the user familiar with the system interaction. The complexity of the interactive exercises and the selection of most suitable modalities are adapted based on user characteristics. The advantage is that users become acquainted with the use of different modalities step by step.

Keywords: Human Computer Interaction, Automatic System Adaptation, Ubiquitous Computing, User Assistance, Task Models.

1 Introduction

In [9], Weiser introduces the paradigm of ubiquitous computing. It describes a world where information is available at any time and any place. Current ubiquitous computing environments are mostly complex systems that are composed of different devices and are connected in heterogeneous networks. The devices are embedded in the surrounding environment and the user should not be aware of them. Frequently, multimodal interfaces are used in order to enable user interaction with different modalities (e.g., voice, touch and keyboard/mouse input). However, while it is already difficult to specify traditional high quality interfaces, the problem in multimodal user interface design is becoming more critical. Especially inexperienced users (e.g., elderly or novel) might be confused with multimodal interfaces since they might have several usability problems. For instance, the user needs to know which tasks are performable and which modalities are available in a certain situation. Additionally, modalities can be combined and each modality can have several commands.

Due to this complexity, we set out to assist the user in order to simplify the handling process of the system.

* Supported by a grant of the German National Research Foundation (DFG), Graduate School 1424, Multimodal Smart Appliance Ensembles for Mobile Applications (MuSAMA).

A. Marcus (Ed.): Design, User Experience, and Usability, Pt I, HCII 2011, LNCS 6769, pp. 706–714, 2011.
© Springer-Verlag Berlin Heidelberg 2011

The proposed assistance can be deployed in lightweight devices (e.g., PDAs or mobile phones) and tries to get inexperienced user familiar with the system interaction. The assistance is able to list performable tasks, to describe interaction modalities and commands and to provide interactive exercises that are adapted according to the level of user experience. Furthermore, for each task the modalities that fit to the user's preference best are selected.

The rest of the paper is structured as follows: In Section 2 we give an overview of related work. Section 3 is concerned with the acquisition of a user profile that includes measurement of user characteristics (experience and preference). Section 4 discusses the idea and specification of smart user assistance for mobile devices. Finally, a summary and future work is presented in the last section.

2 Related Work

In recent years, the generation of assistance in interactive systems has been well recognized. Publications in this area present a vast amount of application fields. Wandke [8] lists amongst other things online help systems in human-computer interfaces, automatic functions in real-time human-machine systems, assistive technology for handicapped people, smart homes, and personal digital assistants. In the following, we aim at presenting some of these systems and compare them with our approach.

Thimbleby and Addison present a system called HyperDoc in [7]. HyperDoc is primarily used for the design and analysis of user interfaces and their manuals. The interactive system as well as its corresponding documentation is specified as finite state machine. Thereby, each state has an associated natural language description and each transition has an associated button name. Due to this specification the system is able to provide answers for user questions. For instance, in order to answer the question “How do I?” the best route from the current state to the required state is determined. By the use of weighted transitions adaptive help can be given. For that reason, weights are adapted based on the current user behavior or previous training with the system. However, finite state machines do not seem to be suitable to model interactive systems. One problem is the large number of states in such systems. Subsequently, simple transition networks are not able to represent the hierarchical structure of an interactive system [3].

There are several assistive systems for handicapped people that apply media adaptation techniques in order to present relevant information with other output modalities. The systems differ essentially by the technique of media adaptation. While a lot of systems are using fixed rules, there are some systems that consider such a process as optimization problem. In the following, the AVANTI [5] system that uses a rule-based media adaptation approach and the CUMAPH [2] system that uses an optimization based approach [1] are presented.

AVANTI [5] is able to display interactive views of adaptive multimedia Web documents. The authors distinguish between adaptability and adaptivity. Adaptability denotes adaptations that take place during the initiation of an interaction (e.g., preferences and abilities), whereby adaptivity denotes adaptations that occur at runtime (e.g., changing user characteristics, situations or familiarity with specific tasks).

AVANTI provides both techniques that allow the system to tailor the user interface according to abilities, skills, requirements and preferences of the user. AVANTI is also capable to support handicapped people (e.g., usage of voice input for visually impaired users). Besides adaptive web documents, traditional web documents can be displayed as well.

Habieb-Mammar et al. propose their system CUMPAPH (Cognitive User Modeling for Adaptive Presentation of Hyper-documents) in [2,6]. CUMAPH is able to adapt hyper documents that are specified as XML-documents. XML-documents consist of several blocks and each block has several elements that can be presented with different media. By computing a compatibility matrix that includes all possible combinations of elements, the system is able to select the most suitable combination of interactive elements. Therefore, the authors have two metrics: the first metric is used to compute the best combination that fit to a cognitive user's profile (generated by a sequence of interactive exercises); the second one favors a combination of media. The combination for which the sum of both metrics is the greatest one is selected.

The problem of media-adaptation is quite similar to ours. However, we focus more on input modality assistance and less on output modality assistance and in contrast to both systems; our main objective is a stepwise familiarization with the abundance of interaction opportunities. Thus, we are not tailoring the system interface itself. Instead, assistance consists of explaining modalities and corresponding commands, and providing adaptable interaction exercises to get the user familiar with the system interaction. The interactive exercises are specified in terms of predefined task models and are substituted according to the provided user information. As a consequence, inexperienced users can become acquainted with the system interaction step by step.

3 User Profile Acquisition

User profiles are employed to store personal characteristics of different users. Based on those user profiles, we are able to adapt the assistance. At the moment, we are measuring two characteristics: *user preference* and *user experience*.

Preference is used in order to determine which modalities have a higher priority since a task can be performed with several input modalities. A stammering user for example, might want to assign voice input a lower priority than people with normal speech. We measure user preference in the following way: if the user handles the mobile application for the first time he/she has to create a profile and adjusts scales with different input modalities. The value ranges of the scales are from zero to ten, whereby one means the lowest priority and ten the highest one. The value zero is used to disable a modality.

Since we need an indicator for determining the complexity of assistance, *user experience* must be regarded as an important characteristic. For example, novel users need assistance for one modality (e.g., write a text with voice input) and basic commands (e.g., start text input or end text input). However, advanced users and experts need assistance for several modalities (e.g., write a text with voice and touch input) and complex commands (e.g., change text properties).

The measurement of user experience takes place as follows: our system stores all modalities and commands in user profiles that are applied by users in order to perform a task. Moreover, if the user is working with the system for the first time, he/she can select well-known modalities and commands. Based on these data, we assign each user into one of the groups shown in Table 1.

Table 1. Categories of Users and their Experience

User Category	Experience
Novel	User knows neither modalities nor commands to perform the task
Experienced	User knows one modality and corresponding basic commands to perform the task
Advanced	User knows one modality and all corresponding commands to perform the task
Expert	User knows all modalities and corresponding commands to perform the task

For example, the task *Write Text* can be performed with voice, touch and keyboard/mouse input. The most suitable combination according to the preference could be: voice input in order to write a first draft and correct possible errors with keyboard/mouse input. Basic voice commands (e.g., start input or end input) are used to begin and start the text input. As soon as the user is familiar with this concept, we adapt the assistance with respect to higher complexity (e.g., several modalities or complex commands).

4 Specification of Smart User Assistance

4.1 Overview

Section 3 presented the acquisition of a user profile. This section describes a method to specify a smart mobile assistance system in ubiquitous computing environments. The main goal of our assistance is a stepwise familiarization of users with several input modalities. Therefore, we are using interactive exercises that are specified in terms of predefined task models and are substituted according to the user profile. Fig.1 depicts a Concur Task Tree (CTT) [4] model that shows how our mobile assistance is given. CTT is a graphical notation to model user activities and distinguishes between different kinds of tasks (e.g., abstract, user, interaction and application). Subsequently, tasks are hierarchical decomposed and linked with each other by temporal operators (e.g., enabling, concurrent etc.) in order to specify the relationship between them. In Fig.1, the abstract task “Select Performable Task” displays all performable tasks according to the context. After that, the user selects the task for which assistance is desired.

The application task “Select Assistance” chooses the most appropriate interaction modality and corresponding interactive exercise to perform the task. Based on this selection the interactive exercise is composed and added dynamically.

Thus, the process of giving assistance can be divided into three parts: determination of performable tasks, selection of most appropriate input modalities and dynamical model composition. Each of these parts is described in the following subsections.

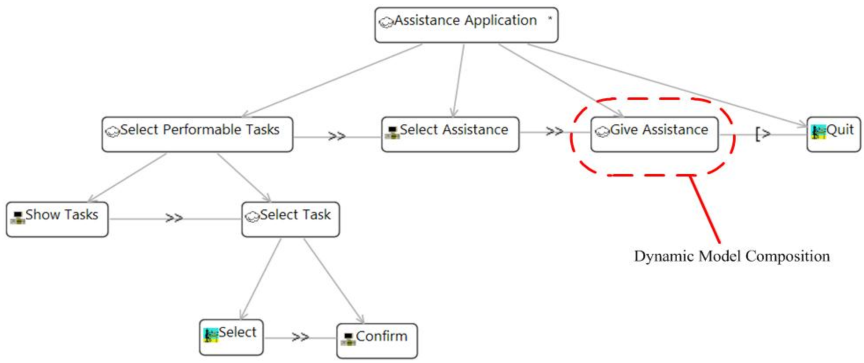


Fig. 1. Specification of Assistance

4.2 Determine Performable Tasks

The process of giving assistance starts with the determination of performable tasks. To accomplish this, we are modeling the ubiquitous computing system with the collaborative task modeling language (CTML) [10]. CTML is a task-based specification language for ubiquitous computing environments. In CTML, user roles are expressed as task-models and the behavior of each role can be expressed by an associated collaborative task expression [10]. By analyzing of these expressions, we are able to determine all performable tasks for one user.

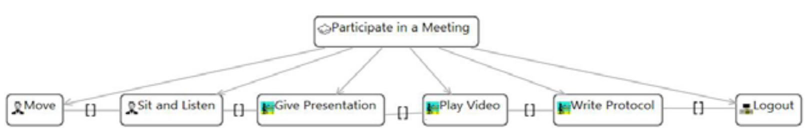


Fig. 2. User Role Participate in a Meeting

The task model in Fig. 2 shows a possible user role for a meeting in a smart meeting room. Possible task constraints (preconditions) could be: the user has to be in the presentation zone to perform the task “Give Presentation”; A DVD has to be inserted into the player to perform the task “Play Video”.

The user has to have left the room to perform the task “Logout”. Thus, if the user is in the presentation zone, the performable tasks are: “Give Presentation” and “Write Protocol”.

If a DVD is in the DVD-Player and the user is in the presentation zone, the performable tasks are: “Give Presentation”, “Write Protocol” and “Play Video”. After determination of all performable tasks, the user chooses the task for which assistance is desired.

4.3 Selection of Input Modalities

After determination of the task, it is necessary to select which modalities should be used to perform the task. There are different solutions to reach this goal. One solution is that the user selects the desired modalities himself/herself. However, this approach is not applicable in each case since some users do not know which input modalities are most suitable. A better solution would be to let the developer decide which modalities should be used to perform a task. Nevertheless, predefined modalities might be unacceptable for handicapped people. Speech-impaired people, for example, are not able to interact via voice input. Consequently, user preferences have to be considered as described in Section 3 and to deselect a modality according to the user profile. For example, if the user deselects speech in the user profile (priority zero); all modality-combinations that include speech-input are deselected. After that, the most suitable modality chosen by the developer is selected unless the user has specified a preference for one of the modalities.

4.4 Dynamic Model Composition

Since our assistance should dynamically adapt interactive exercises according to the user experience, we need for each user category an exercise of different complexity. Table 2 gives an overview what kind of interaction exercise is given for each user category.

Table 2. User Categories and the corresponding Complexity

User Category	Experience
Novel	Interactive exercise for selected modality and corresponding basic commands
Experienced	Interactive exercise for all commands of the selected modality
Advanced	Interactive exercise for new modality and corresponding commands
Expert	Interactive exercise using different modalities

The interactive exercises are specified in terms of predefined task models and substitute the task “Give Assistance”, as illustrated in Fig. 1.
For example, the user asks for assistance for the task “Write Protocol” (Fig. 2).

Available modalities in order to perform the task are the following ones:

- Speech input and correction of possible text errors with mouse/keyboard input
- Stylus input
- Keyboard/mouse input

Additionally, there is the constraint that the modalities speech and stylus input cannot be used concurrently. Let us assume that the stylus input is preferred by the user. It is also assumed that it is a novel user. For this case, a possible interaction exercise is depicted in Fig. 3. At first, the user gets an explanation of the interactive exercise. This includes a general explanation how to use the stylus and basic commands (e.g., open and close a text input). After that, the user has to perform an interactive exercise. This means, that he has to pick up a stylus and to write a text. For writing of a text the user has to open a text input writes something and closes the input.

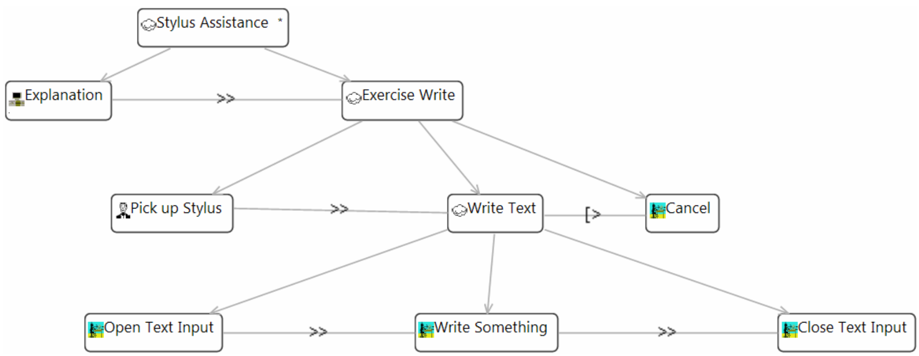


Fig. 3. Specification of Assistance for Novel Users

As soon as the user was able to fulfill the exercise and is familiar with this interaction concept, he is regarded as experienced. Experienced users are able to perform tasks with one modality (the most suitable according to their user profile) and they know all corresponding commands. Fig. 4 shows how further commands are assisted (e.g., change text properties) and how the interactive exercise is adapted.

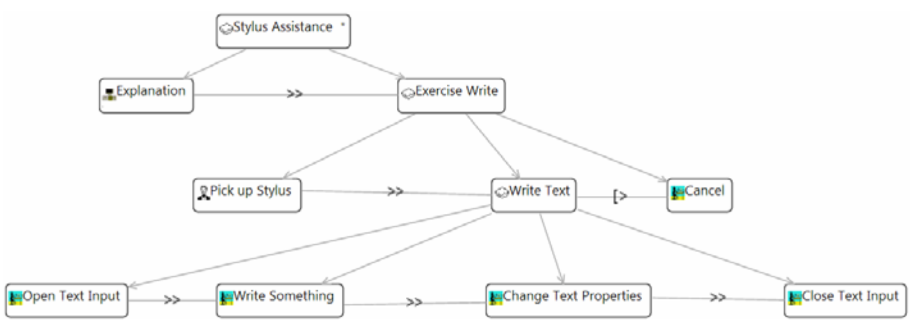


Fig. 4. Specification of Assistance for Experienced Users

This adaptation process is continued with advanced users and experts as well. Advanced users perform the task with other modalities than stylus (e.g., speech input and mouse/keyboard input).

Finally, experts get assistance while using different modalities to perform the task. Fig. 5 and Fig. 6 show the assistance for advanced users and experts.

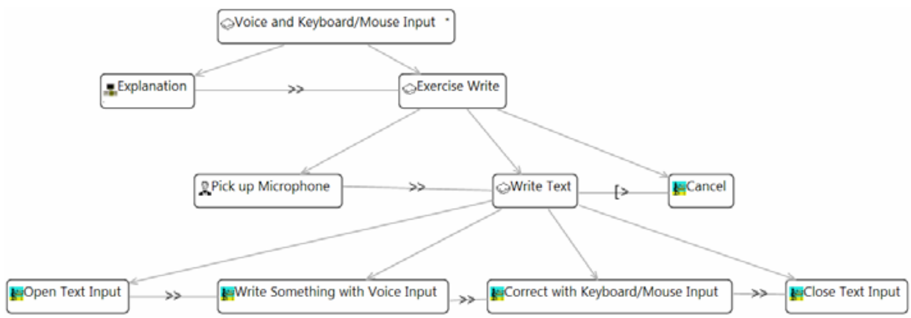


Fig. 5. Specification of Assistance for Advanced Users

To become familiar with voice input that can be corrected with keyboard and mouse there will be an explanation how this can be reached. After the explanation there is again a task a user has to perform. After fulfilling the task the user is expected to have learned the corresponding interaction with the modalities in focus. If there is a problem the exercise can be interrupted and the assistance can start again. This can be seen by looking at the root task that is iterative ("*").

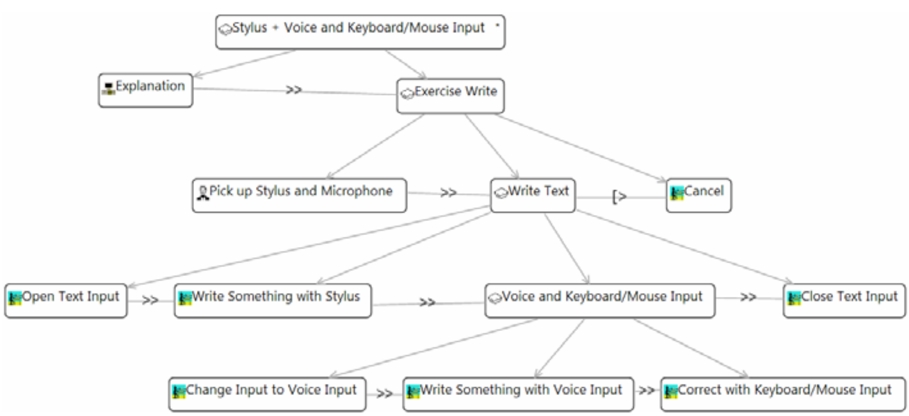


Fig. 6. Specification of Assistance for Experts

Again after an explanation an exercise has to be performed. For this model it is important to note that the hierarchy for the three tasks "Change Input to Voice Input", "Write Something with Voice Input" and "Correct with Keyboard/Mouse Input" was introduce for graphical reasons only. Six tasks at the same level made the model unreadable.

5 Conclusion

The goal of the approach described in this paper is a specification of smart user assistance for mobile devices in ubiquitous computing environments. The assistance provides support in using different input modalities especially for inexperienced user. At first, our assistance determines performable tasks and corresponding input modalities. The modalities are favored according to the given user preference. Thereafter, an interactive exercise is explained and finally performed by the user. The interactive exercises are specified in terms of predefined task models and are substituted according to the provided user experience. As soon as the experience is increasing, the task models are substituted by more complex interactive exercises. The proposed system specification is of particular interest for inexperienced users (e.g., elderly or novel) since they become acquainted with the system step by step.

In the future, we want to evaluate our system based on user studies. We assume that the results support our thesis that changing the underlying models during run time allows a personalized support that is appreciated by users. We assume that interaction with multimodal systems becomes easier with this approach and thus an improvement in the system's usability and accessibility can be reached.

References

1. Bunt, A., Carenini, G., Conati, C.: Adaptive content presentation for the web. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) Adaptive Web 2007. LNCS, vol. 4321, pp. 409–432. Springer, Heidelberg (2007)
2. Habieb-Mammar, H., Tarpin-Bernard, F.: CUMAPH: Cognitive User Modeling for Adaptive Presentation of Hyper-documents. An Experimental Study. In: Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 718–736. Springer, Heidelberg (2004)
3. Kieras, D., Polson, P.G.: An approach to the formal analysis of user complexity. International Journal of Man-Machine Studies 22(4), 365–394 (1985)
4. Paterno, F.: Model-based design of interactive applications. Intelligence 11(4), 26–38 (2000)
5. Stephanidis, C., Paramythis, A., Sfyrakis, M., Stergiou, A., Maou, N., Leventis, A., Paparoulis, G., Karagiannidis, C.: Adaptable and adaptive user interfaces for disabled users in the AVANTI project. In: Campolargo, M., Mullery, A. (eds.) IS&N 1998. LNCS, vol. 1430, pp. 153–166. Springer, Heidelberg (1998)
6. Tarpin-Bernard, F., Habieb-Mammar, H.: Modeling elementary cognitive abilities for adaptive hypermedia presentation. User Modeling and User-Adapted Interaction 15(5), 459–495 (2005)
7. Thimbleby, H., Addison, M.: Intelligent adaptive assistance and its automatic generation. Interacting with Computers 8(1), 51–68 (1996)
8. Wandke, H.: Assistance in human–machine interaction: a conceptual framework and a proposal for a taxonomy. Theoretical Issues in Ergonomics Science 6(2), 129–155 (2005)
9. Weiser, M.: The computer for the 21st century. Scientific American, pp. 94–104 (1991)
10. Wurdel, M., Sinnig, D., Forbrig, P.: CTML: Domain and Task Modeling for Collaborative Environments. Journal of Universal Computer Science 14(19), 3188–3201 (2008)

Author Index