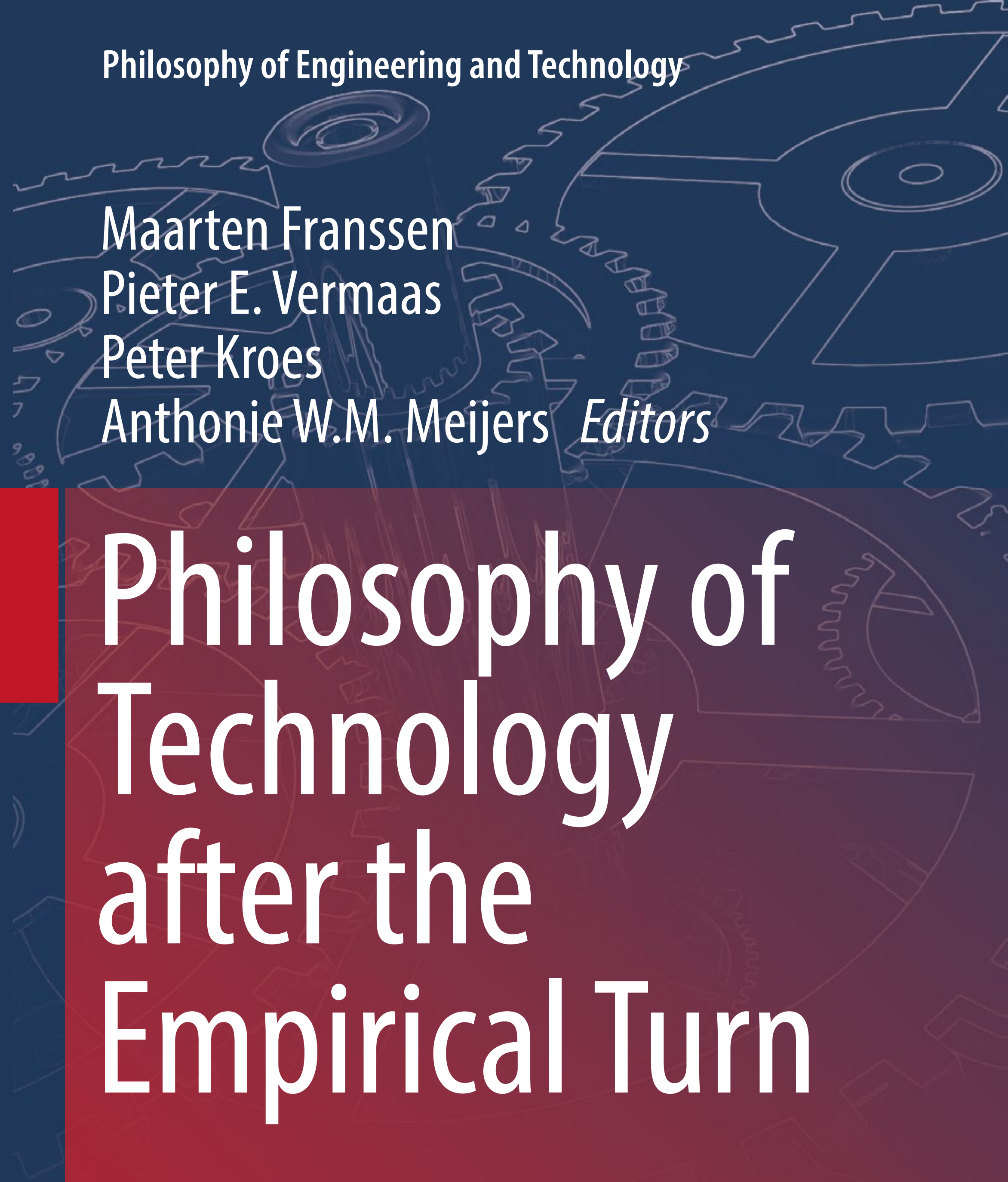


Philosophy of Engineering and Technology

Maarten Franssen
Pieter E. Vermaas
Peter Kroes
Anthonie W.M. Meijers *Editors*

Philosophy of Technology after the Empirical Turn

Springer

Philosophy of Engineering and Technology

Volume 23

Editor-in-chief
Pieter E. Vermaas, Delft University of Technology, The Netherlands
General and overarching topics, design and analytic approaches

Editors
Christelle Didier, Lille Catholic University, France
Engineering ethics and science and technology studies
Craig Hanks, Texas State University, U.S.A.
Continental approaches, pragmatism, environmental philosophy, biotechnology
Byron Newberry, Baylor University, U.S.A.
Philosophy of engineering, engineering ethics and engineering education
Ibo van de Poel, Delft University of Technology, The Netherlands
Ethics of technology and engineering ethics

Editorial advisory board
Philip Brey, Twente University, the Netherlands
Louis Bucciarelli, Massachusetts Institute of Technology, U.S.A.
Michael Davis, Illinois Institute of Technology, U.S.A.
Paul Durbin, University of Delaware, U.S.A.
Andrew Feenberg, Simon Fraser University, Canada
Luciano Floridi, University of Hertfordshire & University of Oxford, U.K.
Jun Fudano, Kanazawa Institute of Technology, Japan
Sven Ove Hansson, Royal Institute of Technology, Sweden
Vincent F. Hendricks, University of Copenhagen, Denmark & Columbia University, U.S.A.
Don Ihde, Stony Brook University, U.S.A.
Billy V. Koen, University of Texas, U.S.A.
Peter Kroes, Delft University of Technology, the Netherlands
Sylvain Lavelle, ICAM-Polytechnicum, France
Michael Lynch, Cornell University, U.S.A.
Anthonie Meijers, Eindhoven University of Technology, the Netherlands
Sir Duncan Michael, Ove Arup Foundation, U.K.
Carl Mitcham, Colorado School of Mines, U.S.A.
Helen Nissenbaum, New York University, U.S.A.
Alfred Nordmann, Technische Universität Darmstadt, Germany
Joseph Pitt, Virginia Tech, U.S.A.
Daniel Sarewitz, Arizona State University, U.S.A.
Jon A. Schmidt, Burns & McDonnell, U.S.A.
Peter Simons, Trinity College Dublin, Ireland
Jeroen van den Hoven, Delft University of Technology, the Netherlands
John Weckert, Charles Sturt University, Australia

The Philosophy of Engineering and Technology book series provides the multifaceted and rapidly growing discipline of philosophy of technology with a central overarching and integrative platform. Specifically it publishes edited volumes and monographs in: the phenomenology, anthropology and socio-politics of technology and engineering the emergent fields of the ontology and epistemology of artifacts, design, knowledge bases, and instrumentation engineering ethics and the ethics of specific technologies ranging from nuclear technologies to the converging nano-, bio-, information and cognitive technologies written from philosophical and practitioners perspectives and authored by philosophers and practitioners. The series also welcomes proposals that bring these fields together or advance philosophy of engineering and technology in other integrative ways. Proposals should include: A short synopsis of the work or the introduction chapter. The proposed Table of Contents The CV of the lead author(s). If available: one sample chapter. We aim to make a first decision within 1 month of submission. In case of a positive first decision the work will be provisionally contracted: the final decision about publication will depend upon the result of the anonymous peer reviewof the completemanuscript.We aim to have the completework peer-reviewed within 3months of submission. The series discourages the submission of manuscripts that contain reprints of previous published material and/or manuscripts that are below 150 pages / 75,000 words. For inquiries and submission of proposals authors can contact the editor-in-chief Pieter Vermaas via: p.e.vermaas@tudelft.nl, or contact one of the associate editors.

More information about this series at http://www.springer.com/series/8657

Maarten Franssen • Pieter E. Vermaas
Peter Kroes • Anthonie W.M. Meijers
Editors

Philosophy of Technology after the Empirical Turn

Springer

Editors
Maarten Franssen
Department of Philosophy
Delft University of Technology
Delft, The Netherlands

Pieter E. Vermaas
Department of Philosophy
Delft University of Technology
Delft, The Netherlands

Peter Kroes
Department of Philosophy
Delft University of Technology
Delft, The Netherlands

Anthonie W.M. Meijers
Department of Philosophy and Ethics
Eindhoven University of Technology
Eindhoven, The Netherlands

ISSN 1879-7202 ISSN 1879-7210 (electronic)
Philosophy of Engineering and Technology
ISBN 978-3-319-33716-6 ISBN 978-3-319-33717-3 (eBook)
DOI 10.1007/978-3-319-33717-3

Library of Congress Control Number: 2016944221

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Contents

Chapter 1
Editorial Introduction: Putting the Empirical Turn into Perspective

Maarten Franssen, Pieter E. Vermaas, Peter Kroes, and Anthonie W.M. Meijers

About 15 years ago, Peter Kroes and Anthonie Meijers published as editors a collection of papers under the title *The empirical turn in the philosophy of technology* (Kroes and Meijers 2000). Next to containing several examples of the kind of studies the editors had in mind, the book made an ardent plea for a reorientation of the community of philosophers of technology toward the *practice* of technology and, more specifically, the *practice* of engineering, and sketched the likely benefits for the field of pursuing the major questions that characterize it in an *empirically* informed way.

This call for an empirical turn, as welcome as it was at the time and as fruitful as it arguably has worked out, was not, of course, an entirely new and audacious beginning. In the broader field of studies dedicated to technology and engineering, the publication of *The social construction of technological systems* (Bijker et al. 1987) had presented technology as a topic meriting serious investigation as a social phenomenon from the perspective of social science and social theory. Under the influence of empirical work produced in this discipline, already during the 1990s several philosophers of technology adopted a less antagonistic and more pragmatic approach to technology (see Brey 2010). This development in the field of social studies of technology was in its turn inspired by the earlier discovery of science as a phenomenon that merited study by the social sciences. That discovery (where we ignore the earlier discovery of science as a norm-guided practice by the sociologist Robert Merton in the 1940s) occurred partially as a result of the huge expansion of science

---

M. Franssen • P.E. Vermaas (✉) • P. Kroes
Department of Philosophy, Delft University of Technology, Delft, The Netherlands
e-mail: m.p.m.franssen@tudelft.nl; p.e.vermaas@tudelft.nl; p.a.kroes@tudelft.nl

A.W.M. Meijers
Department of Philosophy and Ethics, Eindhoven University of Technology,
Eindhoven, The Netherlands
e-mail: a.w.m.meijers@tue.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_1

during the decades following the Second World War and the subsequent attempts by governments to get a grip on this development, which led to an interest in science policy, scientometrics, and similar topics, and the birth of the journal *Social Studies of Science* in 1971. It also partially occurred as a result of developments within the social sciences and the philosophy of science, which cumulated in the proposition of the 'Strong Programme' in the sociology of knowledge by the so-called Edinburgh School led by David Bloor (1976). That programme's move away from the perceived traditional approach of merely 'socially explaining' deviating science toward its aim of socially explaining all of science, was in turn, although independently motivated in sociological terms, considerably facilitated by the upheaval that the work of Thomas Kuhn had caused in the philosophy of science. If anything, it is Kuhn's famous book *The structure of scientific revolutions* (1962) that must be given the credit of having taken the first and audacious step of confronting theories of scientific belief and theory acceptance with the characteristics of living science, and thus initiating the period of empirical turns, not to say empirical turmoil.

Although not conceived as such, Kuhn's book was perceived as criticizing the unrealistic picture of science that underlay the overly abstract and formal philosophy of science that formed the heritage of logical empiricism. To be sure, the empirical turn that Kuhn brought to the philosophy of science was still modest; it consisted mainly of the historical details of a few major turning points in the history of science, such as the Copernican and Galilean revolutions in physics and the chemical revolution instigated by Lavoisier. The response among many philosophers of science to the fierce debates between Kuhnians and Popperians which dominated the 1960s and 1970s was a strong feeling that more detailed knowledge of working science was necessary. Thus emerged the work first of Nancy Cartwright (1983) and Ian Hacking (1983), then of the so-called 'new experimentalists' (e.g. Galison 1987; Mayo 1994), which enriched the philosophy of science with a large fund of empirical studies of 'working science'. Initially, the science observed at work was almost exclusively physics, but this was gradually widened to include empirical studies of the other sciences. This departure from a monolithic philosophy of science to an acknowledgement of and sensitivity to the differences between the various sciences is also occasionally referred to as an empirical turn in the philosophy of science.

The empirical turn heralded by the publication of *The empirical turn in the philosophy of technology* shares some aspects of this latter development in the philosophy of science. Its aim was to steer the philosophical study of technology away from broad abstract reflections on technology as a general phenomenon toward addressing philosophical problems that can be related directly to 'the way technology works' or to 'technology in the making'. In doing so, it focused primarily on the work of engineers. Accordingly, one of its principal messages was to urge a shift of focus to the *design* of technical artefacts, rather than their later career as constituents of use practices. The empirical turn argued for by Kroes and Meijers in 2000 can, therefore, be seen as completing a twofold empirical turn in the philosophy of technology which echoes the preceding similarly twofold turn in the philosophy of science. Adopting characterizations introduced by Brey (2010), the *engineering-oriented* turn of Kroes and Meijers and the earlier *society-oriented* turn that was started off by the arrival of Science and Technology Studies are the complementary aspects of this twofold turn.

To this brief historical sketch some remarks have to be added about similarities and dissimilarities between the developments as they concern science and as they concern technology. To begin with, notwithstanding not only the acknowledgment of the importance of detailed empirical work for the study of science and technology but also its actual implementation in both fields, the philosophical study of science and its study from a social-science perspective seem to have left less marks on each other than have the philosophical and social-scientific forms of studying technology. Still, due to the basic methodological differences between these forms – a philosophical, that is, conceptual and always partly normative orientation in the one, a social-scientific, that is, empirical orientation in the other – the two forms continue to develop at some distance from each other in technology as well, since it has proved difficult to combine them. Second, and reversely, the impact of detailed studies of historical cases seems to have been greater in the philosophy of science than in the philosophy of technology. Excellent historical studies of the development of certain technologies are available for much of the reviewed period (Layton 1974; Constant 1980; Hughes 1983; Vincenti 1990), but their role both in the philosophy of technology prior to the empirical turn and in current philosophy of technology is modest at most, and the same can be said for the use of historical cases and examples in the philosophy of technology (for an exception, see Kroes 1992).

These similarities and dissimilarities and the review of the historical roots of the empirical turn already pose some questions for a research agenda for the coming decades. This volume of essays cannot and does not aim to exhaustively chart this research agenda. What it does aim for is to assess the fruits of the development sketched above, spanning more or less one generation, and among them in particular the developments set in motion by the publication of *The empirical turn in the philosophy of technology*, and to suggest ways in which the next generation could extend these results. It presents brief selections of and reflections on these results and extensions which are already going on now. In this way it aims to contribute to shaping that research agenda. Its contributions address issues that are likely to figure in many suggestions concerning philosophy of technology's research agenda for the near future, such as the question how the relation between philosophy and practice can be developed further. Several contributions contain proposals on how this relation should be developed. The issues addressed concern the philosophical understanding of the practice of engineering and its products, but also the ethical problems caused by the implementation and use of these products and how investigating the process of their creation enriches these discussions. Particularly with respect to the ethical dimension of technology, it is a topic for assessment how the society-oriented and engineering-oriented turns distinguished above have developed side-by-side, and what the balance is of the tensions that result from divergences in the philosophical orientations employed and the synergies that can be expected in view of their shared interests, and how this balance can be improved in the future.

Finally, several of the book's contributions make clear that to compare, as was done above from a historical point of view, the development of the philosophy of technology to ideas and developments in the philosophy of science will continue to be of value for future research. In the philosophy of science, the determination to look more closely at the practice of concrete scientific fields was accompanied by a rejuvenation of the earlier more theoretical orientation of the field, as a result of which the conceptual, theoretical and argumentative framework of science is kept firmly in view. In this way, an impressive literature has been generated that testifies to a fruitful interaction between these two aspects. This leads to the question whether philosophy of technology could profit from a similar approach, or has done enough to promote it, and how much of the corresponding conceptual, theoretical and argumentative framework of engineering and technology it actually has in view and what benefits are to be expected from bringing into play more of it.

During the past decades, then, the field of philosophy of technology has seen radical and ground-braking changes. The black box in which technology had long remained hidden has been opened wide and its contents have become a primary topic for the philosophy of technology to study. With an unprecedented seriousness and determination philosophical research has now started to engage with the *practice* of technology and engineering, with the content of the great variety of knowledge claims to be found there, with the methodology of design and engineering science, and with the moral issues that technology raises for engineers and policy makers from the earliest stages of problem statement and design concept on.

1.1 The Contributions

The contributions to this book fall apart in two different kinds. One kind, making up its first part, follows up on the discussion of the introduction to *The empirical turn in the philosophy of technology* about what the philosophy of technology is all about. It continues the search for the identity of the philosophy of technology by asking what comes *after* the empirical turn. The other kind, in the second part of this book, follows the call for an empirical turn in the philosophy of technology by showing how it may be realized with regard to particular topics. These contributions focus not so much on what comes after the empirical turn, but what happens *in* (implementing) the empirical turn. Together, therefore, these contributions present an overview of the state of the art of an empirically informed philosophy of technology and of various views on the empirical turn as a stepping stone into the future of the philosophy of technology.

The contributions on the identity of the philosophy of technology in the first part groups themselves roughly around two topics, the one addressing primarily the place and role of the philosophy of technology within the philosophical landscape in general and the other the role of the philosophy of technology in contributing to better technologies for society. The opening article, ‘Toward an axiological turn in the philosophy of technology’ by Kroes and Meijers, sets the stage for the ‘identity’ part of the book by in fact addressing both topics. It discusses the issue whether, as a follow-up step after the empirical turn, the philosophy of technology should opt for a normative stance towards its object of study – what they call a normative axiological turn – and by doing so may contribute to better technologies for society. This issue boils down to the question whether the philosophy of technology is a theoretical or practical form of philosophy or both. They argue that taking a normative axiological turn, as is often advocated in recent times, involves major challenges for the philosophy of technology as a philosophical endeavour. Franssen and Koller focus on the place of the philosophy of technology in philosophy in general. In their 'Philosophy of technology as a serious branch of philosophy: the empirical turn as a starting point', they take stock of the situation in contemporary philosophy of technology and what it takes to transform it into a respectable subfield of philosophy. They argue that in the first place the field needs a much greater degree of systematicity in the topics addressed – they propose as main topics the nature of artefacts, design and use – and the answers given to specific questions with regard to these topics. Secondly, the field needs to draw to a greater degree on philosophical expertise acquired and developed in current foundational analytic philosophy, above all metaphysics and the philosophy of language.

In 'Technology as a practical art', Hansson proposes a revitalization of the philosophy of the practical arts and argues that the empirical turn in the philosophy of technology provides an excellent starting point for widening the philosophy of technology to a general philosophy of the practical arts. He proposes a tentative list of topics for a generalized philosophy of the practical arts. This list has a certain preponderance of themes that are close to philosophy of technology in the empirical turn tradition. Whereas Hansson's proposal may be interpreted as a way to embed the philosophy of technology within the wider philosophical landscape, Pitt's proposal is of a much more radical nature. In his 'The future of philosophy: a manifesto' he claims that the future of philosophy in general is the philosophy of technology. He argues that in order to make philosophy a useful feature of the contemporary intellectual scene, we must disengage from minor analytic exercises that have little or no bearing on one another or the world, and try to understand mankind interacting with the world, which in his opinion would be to do philosophy of technology. So the philosophy of technology becomes the mother of all philosophical sub-disciplines.

Next follow two contributions that focus more narrowly on the relation between science and technology and between the philosophy of science and the philosophy of technology. In the field of Science and Technology Studies (STS) it has become quite common to question whether the distinction between science and technology still makes sense in modern times; instead the notion of technoscience has become popular. In 'Science vs. technology: difference or identity? Niiniluoto argues that there is an important conceptual distinction between science and technology. As parts of human culture and society, science and technology exist today in a state of dynamic mutual interaction, but differences can be found in their aims, results, and patterns of development. This means that philosophy of science and philosophy of technology should likewise be in interaction without being reducible to each other. These disciplines have separate agendas which reflect the differences in the aims, results and patterns of development. Nordmann views the relation between science and technology and between the philosophy of science and the philosophy of technology differently. In ‘Changing perspectives: the technological turn in the philosophies of science and technology’, he argues that the philosophy of science and the philosophy of technology both make the same mistake of not taking technology serious enough. The experimental turn in philosophy of science and the empirical turn in philosophy of technology put technology centre stage, yet technology is viewed in both fields mainly through the lens of science, subservient to or derivative of representation and the relation of mind and world. Instead they should look at technology through the lens of working knowledge of the working order of things, not through the lens of things as objects of knowledge but as products of knowledge. He claims that by going through a technological turn both fields will merge in the philosophy of technoscience which will afford a view of research in science and engineering as technological practice.

Nordmann’s paper is the last one of the group that addresses the position of the philosophy of technology in the philosophical landscape. It is followed by a group that focusses on the role of the philosophy of technology in contributing to better technology for society. Here not the identity of the philosophy of technology as a philosophical sub-discipline, but its identity in terms of its relevance for solving societal issues with regard to technology is at stake. In ‘Constructive philosophy of technology and responsible innovation’, Brey argues that the time has come for philosophers of technology to actively engage themselves in the development of responsible technology. He advocates a societal turn, which is a turn from reflective philosophy of technology (academic philosophy concerned with analysis and understanding) to constructive philosophy of technology (philosophy that is directly involved in solving practical problems in society). He describes how a constructive philosophy of technology can contribute to better technology development, better technology policy and better implementation and use of technology, through engineering-oriented, policy-oriented and use-oriented approaches to research. Hillerbrand and Roeser argue for a similar kind of engagement of ethicists in the field of technological risks in their ‘Towards a third ‘practice turn’: an inclusive and empirically informed perspective on risk’. They identify three practice turns in the social and philosophical study of technology that they also relate to risk analysis. The first practice turn singles out technology as a topic meriting serious investigation as a social phenomenon. The second turn steers the philosophy of technology towards the consideration of philosophical problems directly relating to what technology is and what engineers do. The third practice turn explicitly aims at changing the practice of the philosophy of technology by close collaboration with the engineers. Briggle’s paper closes the part that focuses on the societal relevance of the philosophy of technology with a passionate plea for a reconsideration of what the philosophy of technology is all about. In ‘The policy turn in the philosophy of technology’, he critiques the empirical turn for being framed far too much in terms of what philosophers say and not to whom they speak. They talk to their fellow philosophers of technology instead to the ones involved in the actual shaping of technology. He argues for a policy turn which is a turn toward a new model of philosophical research, one that begins with real-world problems as they are debated in public and cashes out its value in real-time with a variety of stakeholders.

The second part of the book contains contributions that in various ways answer to the call for an empirical turn in the philosophy of technology. In ‘A coherentist view on the relation between social acceptance and moral acceptability of technology’, van de Poel explores the implications of the empirical turn for the ethics of technology by investigating the relation between social acceptance (an empirical fact) and moral acceptability (an ethical judgment) of a technology. He develops a coherentist account of the relation between acceptance and acceptability in which empirical facts about social acceptance are related to issues about moral acceptability without assuming that the one entails the other. Houkes’ essay ‘Perovskite philosophy: a branch-formation model of application-oriented science’ illustrates the relevance of detailed case-studies for conceptual issues in the philosophy of science and the philosophy of technology. It uses and develops approaches in these subdisciplines to improve our insight into an on-going phenomenon – the application re-orientation of research – that has so far mostly been addressed in science and technology studies. It answers to the call for an empirical turn by being partly empirical in method: a branch-formation model is developed in discussing a case study of on-going application-oriented research. Zwart and de Vries present and discuss an empirical classification of innovative engineering projects. In their ‘Methodological classification of innovative engineering projects’, they characterize these projects in terms of their overall goal and accompanying method and come up with six different categories. They claim that of these six the engineering means-end knowledge type of projects has been methodologically sorely neglected. They also claim that their empirically grounded categories of types of engineering projects may be a more fruitful starting point for fleshing out the differences between science and technology than by focussing on differences in the kinds of knowledge produced by science and technology. Newberry’s essay ‘For the benefit of humanity: values in micro, meso, macro, and meta levels in engineering’ sketches a four-level taxonomy of values within engineering. In his opinion a close study of the values that inform engineering work may be of crucial importance for understanding the technology-society relationship, since that work is largely proximate to the production of technologies. More specifically, an understanding of the values constitutive of and operating within engineering at a multitude of levels can potentially aid in understanding how engineers go from thought to thing in the processes of design and manufacture. His essay illustrates the importance of empirical data for studying the complex role of values in engineering.

The book closes with three contributions on the notion of (technical) function. This notion has played a key role in the *Dual Nature of Technological Artefacts* research programme which was intended as an illustration of the empirical turn in the philosophy of technology. These papers review and critique the work done on the notion of function from within the Dual Nature program, the paper by Vermaas, and from without, the papers by Feenberg and Schyfter. To begin with the former, Vermaas reviews and critiques the way in which the notion of technical function has been analysed in the Dual Nature program. In ‘An engineering turn in conceptual analysis', he shows that technical function is a term that is intentionally held polysemous in engineering, and argues that conceptual analysis informed by engineering practices should chart and explain this polysemy. The Dual Nature program aimed, however, at determining a single meaning of the term technical function and developed an approach to conceptual analysis, called conceptual engineering, for arriving at this single meaning on the basis of engineering practices. He concludes that this conceptual engineering approach is ill-suited as conceptual analysis of the term technical function in engineering. Nevertheless he considers it to be a useful tool in conceptual analysis, since it can make explicit how specific meanings of polysemous engineering terms are useful to specific engineering tasks. In 'The concept of function in critical theory of technology', Feenberg takes the critical theory of technology as his point of departure for critiquing the analytic philosophy of function as exemplified in the Dual Nature program. He observes that it has made considerable progress in the conceptual analysis of function, but it has not considered the link between function and culture. Any theory of function must situate it in relation to the culture and way of life it serves and he uses the work of Heidegger and Lukács to offer perspectives on how this may be done. Finally, in 'Function and finitism: a sociology of knowledge approach to proper technological function' Schyfter critiques the Dual-Nature-notion of function, in particular the notion of proper function, from the perspective of the sociology of knowledge. He argues against the idea that the notion of proper function precedes the notion of correct use, that is, that correct use may be defined as use that corresponds to an antecedent, fixed proper function. He presents an alternative conceptualisation of technological function. Using finitism, a series of tools developed by the Edinburgh School, he posits that proper functions are socially-endorsed use. In his opinion finitism can serve the 'empirical turn' because it offers analytic tools and methods to clarify the concept of technological function using empirical investigation. Taken together, these three contributions make clear that still a lot of conceptual and empirical work remains to be done in clarifying one of the key notions of modern engineering practice.

Reviewing the various contributions to this book we may conclude that the empirical turn remains a fruitful signpost to follow for the near future of the philosophy of technology. However, we also have to conclude that a shared view on the identity of the philosophy of technology and of its far future is still lacking. The field is still wrestling with both its philosophical profile and its societal relevance. On the one hand that may be a situation to be deplored, on the other hand that situation gives rise to interesting and vigorous discussions about what the philosophy of technology is all about, about how best to approach technology as a topic for philosophical reflection and about how to implement these approaches. Such discussions are and ought to be an integral part of any discipline on its road to maturity.

References

Bijker, W. E., Hughes, T. P., & Pinch, T. (Eds.). (1987). *The social construction of technological systems: New directions in the sociology and history of technology*. Cambridge, MA: MIT Press.

Bloor, D. (1976). *Knowledge and social imagery*. London: Routledge & Kegan Paul.

Brey, P. (2010). Philosophy of technology after the empirical turn. *Technè, 14*(1), 36–48.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Oxford University Press.

Constant, E. W. (1980). *The origins of the turbojet revolution*. Baltimore: Johns Hopkins University Press.

Galison, P. (1987). *How experiments end*. Chicago: University of Chicago Press.

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.

Hughes, T. P. (1983). *Networks of power: Electrification in western society 1880–1930*. Baltimore: Johns Hopkins University Press.

Kroes, P. (1992). On the role of design in engineering theories: Pambour's theory of the steam engine. In P. Kroes & M. Bakker (Eds.), *Technological development and science in the industrial age* (pp. 69–98). Dordrecht: Kluwer Academic Publishers.

Kroes, P., & Meijers, A. (Eds.). (2000). *The empirical turn in the philosophy of technology*. Amsterdam: JAI-Elsevier.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Layton, E. (1974). Technology as knowledge. *Technology and Culture, 15*(1), 31–41.

Mayo, D. G. (1994). The new experimentalism, topical hypotheses, and learning from error. In D. M. Hull, M. Forbes, & R. Burian (Eds.), *PSA 1994* (Vol. 1, pp. 270–279). East Lansing: Philosophy of Science Association.

Vincenti, W. G. (1990). *What engineers know and how they know it: Analytical studies from aeronautical history*. Baltimore: Johns Hopkins University Press.

Maarten Franssen M.S. in Theoretical Physics, M.A. in History and Ph.D. in Philosophy, all from the University of Amsterdam. Currently Associate Professor with the Philosophy Department of Delft University of Technology, the Netherlands, and engaging with foundational and methodological issues in the philosophy of science and technology. His specific research interests are the nature of normativity in relation to artefacts and their use, the metaphysics of artefacts, instrumental and sociotechnical systems as constitutive of technology, and the conception of design as decision-making and its problems.

Pieter E. Vermaas M.S. in Theoretical Physics from University of Amsterdam, and Ph.D. in Philosophy of Physics from University Utrecht. Currently Associate Professor with the Philosophy Department of Delft University of Technology, the Netherlands. Research interests: the concept of technical function; technical artefacts in engineering and metaphysics; design methodology; design for values. Editing for the *NanoEthics and Philosophy and Technology* journals. Editor-in-Chief of the Springer *Philosophy of Engineering and Technology* book series.

Peter Kroes Engineering degree in physical engineering from the University of Technology Eindhoven and Ph.D. in the philosophy of physics from University of Nijmegen (The Netherlands). He is professor in philosophy of technology at Delft University of Technology. His main areas of interest are philosophy of technology and philosophy of science. Recent book publications: *Artefact kinds; ontology and the human-made world*, eds. Maarten Franssen, Peter Kroes, Thomas A.C. Reydon, Pieter Vermaas, Synthese Library 365, Springer, 2013, *Technical artefacts: creations of mind and matter*, Springer, 2012.

Anthonie Meijers Studied mechanical engineering at Delft University of Technology and philosophy at Utrecht University. He has a Ph.D. in philosophy from Leiden University. He is full professor at the Department of Philosophy and Ethics of Eindhoven University of Technology. His main interests in the field of the philosophy of technology are the theory of artefacts, agency and artefacts, technology and behavioral change, and the epistemology of technology. He is one of the co-editors of the 2000 book on the empirical turn in the philosophy of technology and the editor in chief of the *Handbook Philosophy of Technology and Engineering Sciences* (2009). He is also editor and co-founder of Philosophical Explorations, an international journal for the philosophy of mind and action.

Chapter 2
Toward an Axiological Turn in the Philosophy of Technology

Peter Kroes and Anthonie W.M. Meijers

2.1 Introduction: A Brief Look Back

In our introduction to the book *The Empirical Turn in the Philosophy of Technology* we argued for a triple reorientation in mainstream philosophy of technology, namely (1) from a focus on the use of technology and its societal effects to the development of technology, in particular engineering design, (2) from a normative to a descriptive approach and (3) from moral to non-moral issues (Kroes and Meijers 2000). Our main reasons for arguing for this triple reorientation, referred to as "an empirical turn", was the treatment of technology as a black box and the dominance of (negative) normative starting points underlying many of the most influential analyses of technology in the philosophy of technology. We believed (and still believe) that "a better understanding of technology resulting from an empirical turn will contribute to better normative analyses and evaluations" (ibid, p. xxxiii). In this paper we analyze what a turn to better normative analyses and evaluations will imply for the philosophy of technology. We refer to such a turn as an axiological turn in the philosophy of technology. We distinguish between a descriptive and normative axiological turn. The former is very much in line with the empirical turn and the latter deviates from it by trying to reintroduce, in a specific way, a normative element in the philosophy of technology. Our analysis of what is involved in an axiological turn is not to be understood in the sense that we think that the empirical turn in the philosophy of technology has been completed and that now the time has come to turn to normative issues.

P. Kroes (✉)
Department of Philosophy, Delft University of Technology, Delft, The Netherlands
e-mail: p.a.kroes@tudelft.nl

A.W.M. Meijers
Department of Philosophy and Ethics, Eindhoven University of Technology,
Eindhoven, The Netherlands
e-mail: a.w.m.meijers@tue.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_2

Neither do we intend this axiological turn as a return to the old-style normative evaluations of technology as a whole. On the contrary, a descriptive axiological turn is a straightforward implementation of the empirical turn. We argue for a descriptive analysis of the role of various values that play a role in the design, development and use of technology. We see an understanding of the role of values in shaping technologies as a preliminary step to critically evaluate this role and possibly to normatively intervene in the process of developing technology. This last step, a normative axiological turn, we argue not only requires insight into the actual role of values in technology but also a critical self-reflection on the role of philosophers of technology as (possible) actors involved in (developing) technology.

So the aim of this paper is primarily to look ahead – What are the next steps to be taken in the philosophy of technology? – not to look back on what has happened since our plea for an empirical turn was launched. Nevertheless it is important to pause for a brief moment on what has been achieved, and what not, and what lessons were learned. Whether or not our call for an empirical turn and our own attempts to bring it into practice were successful we leave for others to decide. From our point of view one of the most important results of the call for an empirical turn is without doubt the handbook *Philosophy of Technology and Engineering Sciences* edited by Meijers (2009). Furthermore, the project *The Dual Nature of Technical Artifacts* was intended as an exemplar of such a turn and the project did attract a lot of attention within the field. The outcome of our efforts so far has strengthened our conviction that the design and development phase of technology not only is an important domain for fruitful philosophical research, but also that a better philosophical understanding of technology "itself" may shed new light on long-standing issues in traditional philosophy of technology, for instance on the moral status of technical artifacts (see Kroes 2012). Of course, much work remains to be done, but we still believe that a shift in focus along the lines indicated above is very promising and needed for a better understanding of technology and how it affects or even defines the modern human condition.

Two lessons-learned are worth pointing out here. The first one concerns the connection between philosophy of technology and mainstream philosophy. Although the philosophy of technology has shown a rapid expansion during the last two decades and has strengthened its institutional base,[1] the philosophy of technology still is a marginal field within philosophy. In our opinion it would be fruitful to strengthen its ties with mainstream philosophy, in order to learn from its results and to draw attention to the fundamental philosophical problems raised by modern technology. Given that the modern human condition is more or less defined by technology, the virtual absence of technology as a topic of philosophical reflection of its own in mainstream philosophy is puzzling, to say the least.[2] In a special issue of

---

[1] For instance, a new journal dedicated to the philosophy of technology (*Philosophy and Technology*, edited by Floridi) and a new book series (*Philosophy of Engineering and Technology*, edited by Vermaas) have been set up.

[2] In their contributions to this book, Pitt and Nordmann argue for a much more central place of philosophy of technology in mainstream philosophy.

Studies in History and Philosophy of Science devoted to the Dual Nature project we purposely involved 'mainstream' philosophers by letting them comment on results from this project (Kroes and Meijers 2006).

We are fully aware that apart from its benefits, a more central place of the philosophy of technology in mainstream philosophy may have its own drawbacks, especially from the point of view of those who argue for an active role of the philosophy of technology in shaping (the role of) modern technology in society. In his article *Does Philosophy Matter*? Borgmann (1995, p. 295) argues for a negative answer to this question; he claims that "By most any measure of social or cultural prominence, academic [i.e. mainstream] philosophy does not matter in contemporary life." One of his arguments is that philosophers have locked themselves up in their ivory tower by primarily writing for other philosophers, without any groups outside their ivory tower picking up on their results. Of course, Borgmann's assessment of the impact of philosophy in general on modern life may be disputed (for example for applied ethics), but for many subfields in mainstream philosophy it indeed appears to be the case that philosophers write primarily only for their colleagues. In that sense philosophy is not different from disciplines such as theoretical physics, much of mathematics, and history. By moving the philosophy of technology in a mainstream direction it may run the risk of having to face the same charge of irrelevance for contemporary life by drawing attention away from the pressing moral issues of modern technology. After all, the empirical turn is first and foremost a call for understanding technology itself. So, why would this understanding matter? In our opinion, the situation with regard to (understanding) technology may be different from many of the other topics discussed in mainstream philosophy, because, as we remarked above, technology is such a defining feature of the modern condition. Thus, understanding technology may contribute to the intrinsic aim of understanding what it means to be human in contemporary times. Apart from this, however, we think there is another reason why the call for an empirical turn does not enhance the risk of irrelevance. In our call for an empirical turn we have stressed that the underlying *rationale* of this call is that a better understanding of modern technology itself generally speaking is a *conditio sine qua non* for better dealing with moral issues about modern technology. In this paper we therefore set ourselves the task to analyze what the implications are for the philosophy of technology if we take this *rationale* for the empirical turn seriously.

The second lesson concerns the connection between the philosophy of technology and the engineering world. The empirical turn calls for an empirically informed philosophy of technology. Therefore, from the start of the Dual Nature project one of our aims has been to involve engineers in our philosophical endeavor and to engage in an ongoing (critical) dialogue with them. In spite of our efforts and of incidental successes[3] it turns out to be difficult to bring these two worlds together in

---

[3] One of the promising developments is the fPET, the *Forum on Philosophy*, *Engineering and Technology*; its mission is "to foster scholarship and reflection, by scholars and practitioners from diverse fields, including engineers, philosophers, and social scientists, on the topics of engineers, engineering, and technology." See http://web.ecs.baylor.edu/faculty/newberry/fpet-2014.html.

a structural and fruitful way. Herein lies one of the great challenges for the philosophy of technology. It is of crucial importance for avoiding the danger that the philosophy of technology "does not matter", whether or not as part of mainstream philosophy. If there is one field in which a dialogue and cooperation with engineers is urgent it is in the field of ethics and technology. Recently, Mitcham (2014) has voiced a wake-up call to all of us, but to engineers in particular. Referring to Jasper's notion of Axial Age, in which leading intellectuals did not simply accept and started to critically assess the cultures into which they were born, Mitcham maintains that we are entering into a second Axial Age in which the physical and technological world in which we are born is not simply accepted and has to be critically assessed. Engineers have to face the "challenge of thinking about what we are doing as we turn the world into an artifact and the appropriate limitations of this engineering power", but according to Mitcham they lack the means to do so and he advises them to turn to the humanities and social sciences. An axiological turn as discussed in this paper may be seen as a response to Mitcham's wake-up call from the point of view of the philosophy of technology. Just as in the case of the empirical turn, its implementation will require a close cooperation between engineers and philosophers of technology.

The paper is composed in the following way. In order to clarify in what sense the axiological turn is a continuation of the empirical turn and in what sense it is not, we will first have a look at the role of values and norms in the empirical turn (Sect. 2.2). Then we turn to the various kinds of values and norms that play a role in the object of study; for the purpose of this paper we will not focus on technology in general but on engineering practice. The study of these values and norms from an empirical turn perspective leads to what we call a *descriptive axiological* turn (Sect. 2.3). In Sect. 2.4 we discuss whether a purely descriptive axiological turn is feasible. The real break with the empirical turn occurs when non-epistemic values and norms, in particular ethical ones, enter at the meta-level of the approach (analytical framework) of the philosophy of technology. This leads to what we call a *normative axiological* turn (Sect. 2.5). The final Sect. 2.6 analyzes how recent developments in the philosophy of technology relate to our notion of an axiological turn and discusses what challenges will have to be faced when implementing a normative axiological turn.

2.2 Values and Norms of the Empirical Turn

Kroes and Meijers (2000, p. xxii) argued that an empirical turn in the philosophy of technology implies a reorientation with regard to the role of normativity in its approach and in the topics it studies. This does not imply that a philosophy of technology that answers to the call of the empirical turn is in itself value-free. It is still guided, of course, by the values and norms of philosophical analysis, in particular of an empirically informed philosophy of technology. An elaborate discussion of what these norms and values are falls outside the scope of this paper. For our purpose in this paper the following remarks should be sufficient.

The empirical turn calls for an empirically informed philosophy of technology. This brings into play the epistemic value of empirical adequacy: the empirical claims taken into account in philosophical analysis have to be empirically adequate.[4] But how does this value touch upon philosophical claims? Philosophy is generally considered not to be an empirical science; in so far it makes claims that are subject to the norm of empirical adequacy, these claims are not philosophical claims, but empirical ones. So what role can the value of empirical adequacy play in philosophy of technology? Referring to Quine's criticism of the analytic-synthetic discussion, we argued in Kroes and Meijers (2000) that empirical adequacy may play a role because there is no sharp distinction between the empirical sciences and philosophy. Empirical adequacy is a constraint on a whole network of claims that cannot neatly be split into empirical and philosophical claims. So empirical adequacy as a value does constrain philosophical claims but in a rather indirect, 'holistic' way.

Quine's criticism of the analytic-synthetic distinction may throw an interesting light on the question of the role of the value of empirical adequacy in philosophical analysis, but that still leaves the question open which other values and norms play a role. More in particular, we may ask which values and norms are at work in Quine's analysis of the analytic-synthetic distinction, apart from, in whatever way, empirical adequacy.[5] The *kind* of values and norms involved appear to be of an epistemic nature, since Quine makes a claim about the "totality of our so-called knowledge or beliefs" (Quine 1951, p. 42). Thus, values like truth, simplicity, consistency and explanatory power may be of relevance. A value of particular importance for philosophy is, in our opinion, conceptual coherence: philosophical views and claims are evaluated on their internal coherence and on their external coherence, that is, with the rest of the whole fabric of (knowledge) claims. The importance of the value of coherence for philosophical thought is illustrated by the coherence theories of truth which try to explicate the fundamental value of truth is terms of the value of coherence.

The notion of coherence offers an interesting way to connect the value of empirical adequacy to philosophical work without drawing philosophy in the domain of the empirical sciences. Philosophical views and claims have to be coherent with well-established empirical claims. What kind of constraints this imposes on philosophy depends of course on how the notion of coherence is explicated. In this respect it is rather surprising to note that the notion of coherence itself is seldom explicitly the topic of philosophical analysis (see Kroes 2006).

Given the above remarks we will assume in the following that the values and norms that play a role in a philosophical analysis of engineering practice that

---

[4] We will not enter here in a discussion whether empirical adequacy is a value, a norm or both. If empirical adequacy is taken to be a norm, then truth may be taken to be its corresponding value.

[5] Quine makes a philosophical claim about the analytic-synthetic distinction and this claim itself is, just as other philosophical claims, constrained by empirical adequacy (by experience) only at the "edges" of the fabric of our beliefs.

answers to the call of the empirical turn are of an epistemological nature. Let's now turn to the values and norms that play a role in engineering practice itself.

2.3 A Descriptive Axiological Turn: Values and Norms in Engineering Practice

A common way of looking at engineering practices is that they are embedded in broader processes that aim at the production of valuable goods and services.[6] In what respect a produced good is valuable depends on the stakeholders involved. Design engineers may highlight the technical value by stressing technical innovations in and patents on the product, or the potential value for users, whereas production managers may look at the value created primarily in terms of corporate profits, and sales managers in terms of market position. The end users may appreciate the value of the goods and services in terms of satisfying their needs and reaching their goals; these needs and goals may be very diverse bringing into play various kinds of user values. Governmental institutions may look at how the creation, production and use of technical goods and services enhance public or social values like the health and safety of production workers or users or the privacy of citizens. Although these various values are associated with different phases and stakeholders in the product creation process, engineers involved in this process will have to take these values into account in their work. Nowadays, values related to health, safety and sustainability, for instance, play a central role in engineering design practice, apart from technical and economic values.

So, engineering practice is a thoroughly value-laden, normative practice. In view of this we advocate a descriptive axiological turn in the philosophy of technology, which is the empirical and philosophical study of engineering practices as value-laden, normative practices.[7] Following MacIntyre's definition, values ("ends and goods") and norms ("standards of excellence") play a key role in any practice (MacIntyre 1984, p. 187):

> By a 'practice' I am going to mean any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended.

What makes engineering practice, compared to many other practices, so interesting is that many different kinds of values play a role in it. This raises many questions. What different kinds of values are involved in engineering practice? Is there a hierarchy among these different kinds? How do engineers conceive of or define

---

[6] This paragraph is based on (Kroes and van de Poel 2015).

[7] Our call for a descriptive axiological turn may be seen as a call to put off the self-imposed narrow blinkers of the empirical turn with its strict focus on technical values.

values? How do they operationalize them? How do engineers deal with these various values? How do they handle conflicts or trade-offs between different values? Answering these questions requires not only empirical but also philosophical research. What is needed is a conceptualization or a theory of value that helps in explicating the meaning of the notion of value in the above characterization of engineering as a value laden practice. Furthermore, of particular interest from a philosophical point of view is the fact that engineers have to deal with problems that involve at the same time values and norms that are traditionally associated with different domains of philosophy: epistemic, practical, moral and aesthetic values and norms. From a value point of view engineering practice with its "standards of excellence" does not fit into one of these neatly defined philosophical domains. For a better understanding of the "ends and goods" and "standards of excellence" much empirical and philosophical work remains to be done. The discussion about the moral status of technical artifacts illustrates that a descriptive axiological turn along these lines will indeed involve fundamental philosophical issues (see Meijers 2009, part V; Kroes and Verbeek 2014). One way, namely, to interpret the goods produced by engineering practice is in terms of the technical artifacts that are designed and produced. Whether or not these goods, interpreted in this way, embody moral values or not, has been and still is hotly debated within the philosophy of technology.

Our plea for a descriptive axiological turn in the philosophy of technology is in line with our previous call for an empirical turn in the sense that it requires a descriptive approach, not a normative one. It is not in line in so far as it does draw values other than technical ones more centrally into the object of philosophical analysis, including moral values. In our opinion this is not so much a revision of our original position as well as the obvious next step to be taken.[8] The idea of the empirical turn was and still is to "open the black box of technology" and to draw philosophical attention to the technical (engineering) aspects of the creation of technical artifacts. But from the beginning the idea has also been that "philosophical reflection should be based on empirically adequate descriptions reflecting the richness and complexity of modern technology" (Kroes and Meijers 2000, p. xix). In order to do justice to the richness and complexity a focus on technical aspects, values and norms (such as efficiency, effectiveness, reliability etc.) is not sufficient. It is also necessary to include all the various kinds of value that play a role in engineering practice, including the moral ones.

2.4 Is a Purely Descriptive Axiological Turn Feasible?

A descriptive axiological turn, as a follow-up of the empirical turn, may be necessary for a full understanding of the various values at play in engineering practice, but we have to ask ourselves to what extent it is a feasible option for the philosophy

---

[8] Our plea for a descriptive axiological turn is, as we remarked before, not motivated by an assessment that the empirical turn with its focus on non-moral aspects has been accomplished.

of technology. Is it possible to take a purely descriptive stance towards engineering practice? The answer to this question depends on whether as a matter of principle or as a matter of fact, a sharp distinction between descriptive and normative claims can be maintained. Whereas in our empirical turn paper the distinction between the descriptive (synthetic) and conceptual (analytic) played a crucial role, now the distinction between the descriptive and normative is at stake. This distinction has been challenged in the debate about thin and thick ethical concepts.

The notion of a thick ethical concept was introduced by Williams (1985) in his critique of fact-value theorists who "are bringing their distinction to language rather than finding it there and, in addition, are unreasonably expecting that when the distinction is revealed it will be found very near the surface of language" (p. 130). Williams observes that many concepts, such as treachery, promise, brutality and courage, are neither just descriptive nor just prescriptive/evaluative; their application is determined by a combination of fact and value. He calls them 'thick' ethical concepts. Statements using thick ethical concepts do not fit neatly into the pigeon holes of factual (descriptive) and evaluative (prescriptive) statements. They have features of both. Somebody may, given certain circumstances, be called rightly or wrongly courageous but this expresses a value judgment in addition to a factual statement. When people disagree about whether somebody behaved courageously or not, they may try to resolve their disagreement by analyzing more closely the facts of the matter, that is, the person's behavior under the given circumstances. But they may also try to resolve their disagreement by analyzing and comparing more closely their normative standards for courageous behavior in this case. With regard to the use of thick concepts like courageous, however, there is no guarantee that a recourse to either the facts only or to the normative standards only will be able to resolve the disagreement. According to Williams the application of thick concepts "is at the same time world-guided and action-guiding" (p. 141); they are both descriptive and prescriptive. In contrast to thick ethical concepts, thin ethical concepts, such as good or right, lack any or almost all factual content.[9]

The problem with regard to thick ethical concepts is how to interpret the role of facts and values in determining their meaning. Williams questions a particular account according to which the application of a thick concept is only determined by its descriptive elements; so it assumes that evaluative elements play no role in this. In other words, for any thick concept "you could produce another that picked out just the same features of the world but worked simply as a descriptive concept, lacking any prescriptive or evaluative force" (p. 141). According to this account a thick ethical concept is, therefore, simply a descriptive concept on top of which an evaluative element is added. According to this so-called strong separationist line of thought it is possible to disentangle the facts and values involved in thick concepts.

Williams doubts that it is always possible to come up with a descriptive concept that captures the descriptive content of a thick concept. In the example above, about the disagreement about calling somebody courageous, it may be assumed more or

---

[9] Williams himself does not use the notion of a thin ethical concept.

less tacitly that it would be possible to separate the relevant facts involved in calling somebody courageous. If that would be possible then indeed there are in principle two ways to interpret the disagreement: either there is disagreement about the relevant facts involved or about the moral norms involved (or about both). On a non-separationist account the disagreement is of a different nature; it concerns a unitary whole of facts and values that cannot be disentangled into its factual and evaluative constituents.

How to interpret thick concepts and whether or not they are a kind of concepts *sui generis*, different from factual and evaluative concepts, has become a matter of debate (Kirchin 2013). For our purposes it will not be necessary to enter into this debate. More important for us is that as a matter of fact thick concepts, as Williams suggests, are part and parcel of the surface structure of ordinary language and what this implies for the possibility of a purely descriptive axiological turn. First we will have a brief look at the use of thick concepts in engineering practice.

Given that engineering practice is a thoroughly value-laden practice it comes as no surprise to observe that it is so to speak loaded with thick concepts. Key concepts such as safe, dangerous, efficient, wasteful, reliable, user-friendly, environment-friendly, sustainable, flexible etc. all appear to be thick concepts.[10] On the assumption that these concepts have been operationalized into clear, objective measurable lists of specifications, these concepts may be taken, *prima facie*, to lead to factual statements: “X is safe to use for doing Y”, for instance, then may amount to the factual claim that X satisfies a specific list of (measurable) criteria when used for doing Y. Nevertheless, even in those circumstances, there appears to be an evaluative element involved in the claim, namely an (implicit) recommendation to use X when one wants to do Y. Moreover, the acceptance of the operationalization itself implies a value judgment to the effect that this operationalization of the notion of safety is (morally) acceptable. Thus, although the application of thick engineering concepts in particular circumstances is clearly world-guided, it involves at the same time value judgments (i.e. is world-guiding), even if these concepts have been operationalized in terms of objective measurable criteria.

What implications may be drawn for our distinction between a descriptive and normative axiological turn from Williams’ distinction between thick and thin concepts and from the fact that thick concepts are ubiquitous in engineering practices? To begin with the latter, the fact that the use of thick concepts is part and parcel of engineering practice is not of direct relevance to the distinction between the two axiological turns. It is a highly significant fact about the object of study of each axiological turn but does not, as such, undermine the distinction between these two turns. From the point of view of a descriptive axiological turn the use of thick concepts in engineering practice is an interesting and highly important topic for further empirical study and from a normative axiological point of view the justification of

---

[10] Many more items can be added to this list. For instance, most of the “ilities” that play a key role in software engineering, and outside that field, appear to be thick concepts; see, for instance, http://codesqueeze.com/the-7-software-ilities-you-need-to-know/.

the (often implicit) value judgments contained in or implied by the use of thick engineering concepts calls for a critical analysis.

So, if Williams' view on thick concepts does affect our distinction between a descriptive and normative axiological turn, then it must be at the meta-level, the level of their approaches, not of their object of study. The way we defined both approaches assumes a fact-value distinction, or more precisely a distinction between a fact-oriented epistemic and a normative analytic (conceptual) framework. The approach of a descriptive axiological turn is defined in terms of an analytic framework in which only epistemic values are at work, whereas in the analytic framework of the normative axiological turn also other (moral, practical, aesthetic) values play a role. The concepts of the basic values involved in these frameworks, among which we find the concepts of truth, moral goodness, instrumental goodness and beauty, all appear to be thin concepts. In order for these concepts to be applied in analyzing concrete situations they will have to be explicated in more specific concepts that in turn will have to be operationalized. The concept of truth, for instance, may be explicated in more specific concepts such as empirical adequacy, explanatory power, simplicity, coherence etc. and moral goodness in beneficence, doing no harm, pleasure, pain, utility etc. Now the crucial question is whether the application of some of the more specific concepts in a descriptive axiological turn does or does not, as a matter of principle or as a matter of fact, involve non-epistemic value judgments. If they do, then some of these concepts are thick concepts which undermines the idea that a purely descriptive axiological turn is possible.

We will not pursue the question whether a purely descriptive axiological turn is feasible any further, except for the remark that we have to be careful not to bring the fact-value distinction to our analytical frameworks, instead of, in line with Williams' quote above, "finding it there".

The issue we have been discussing is, of course, closely related to the longstanding debate about whether science is, or may be in principle, value-free.[11] Douglas (2000), for instance, argues that when it comes to setting standards of statistical significance in many parts of science the inductive risks involved may have non-epistemic consequences and that therefore those parts of science are not 'value-free'. Her arguments also apply to setting standards of (statistical) significance in engineering practices. At the end of her paper she concludes (p. 578):

> Finally, there are cases where the science will likely be useful but the potential consequences of error may be difficult to foresee. This gray area would have to be debated case by case, but the fact that such a gray area exists does not negate the basic argument: when non-epistemic consequences of error can be foreseen, non-epistemic values are a necessary part of scientific reasoning.

If we replace science in this quote by philosophy of technology, and if the philosophy of technology claims to be somehow useful in shaping technology – even if this "somehow" involves much deeper shades of gray than in the case of science –

---

[11] This debate goes back at least to Rudner (1953) and the reply by Jeffrey (1956); for a recent contribution to this debate, see Douglas (2000).

then we see no reason why the same should not be the case for the philosophy of technology. This brings us to the normative axiological turn.

2.5 A Normative Axiological Turn: Values and Norms in Philosophy of Technology

A normative axiological turn implies a departure from the empirical turn. Now there is no attempt to stay safely within the realm of value-free inquiry (except, of course, for epistemic values). This turn involves actively taking up normative issues in the philosophical analysis of the engineering practice. There are at least two options for a normative axiological turn:

- The reflective position. According to this view it is the task of the philosopher of technology to analyze normative issues related to technology and engineering practice and to actively participate in societal discussions by preparing and facilitating debates and decisions about technology;
- The substantive position. In this view the philosopher of technology not only analyses normative issues but also takes a normative stance him/herself with regard to the issues at stake and acts accordingly.

It is not our intention to enter here into a discussion of whether these two variants of a normative axiological turn can always be clearly distinguished and which option the philosophy of technology, if any, should choose. Instead we will make a number of observations that are meant to elaborate what we have in mind with a normative axiological turn.

The first thing to mention here is that the second position clearly makes more normative commitments than the first. The first position, however, is itself not free from normative commitments, since there is no value-free analysis and participation in societal debates: issues are framed in a certain way, there is a choice of aspects that are taken into account, some values (debates) are considered to be more important than others, etc. But these normative commitments play a role in the background and are about the way the normative issues at stake should be analyzed, debated and decided upon.

The second thing to observe is that taking a normative axiological turn does not necessarily imply taking a *moral* turn. The normative axiological turn may pertain to all kinds of values and norms: epistemic, moral, practical, aesthetic etc. As we noted in our empirical turn paper, there has been a long normative tradition in the philosophy of science of a non-moral nature. In this tradition various schools within the philosophy of science have criticized the epistemic values and methodological norms applied in actual science on the basis of the epistemic values and methodological norms accepted by those schools. One can easily imagine something analogous to happen in the philosophy of technology. A philosopher of technology may criticize on the basis of her/his epistemic values and methodological norms a particular engineering practice for believing that an engineering theory is sufficiently supported. In that case there is a clash between two different sets of epistemic values and norms. Similarly, starting from some set of values and norms for practical, instrumental action a philosopher of technology may criticize the use of a particular theory as a reliable guide for action. And, of course, by taking a *moral* stance a philosopher of technology may criticize moral principles of and decisions taken by engineers. By taking a moral stance not only principles and decisions that are moral in nature, but also morally *relevant* claims and decisions may be criticized. Within a particular context of action, believing a theory to be sufficiently supported for accepting that theory as a reliable basis for action may be highly morally relevant. In other words, taking a moral normative stance may involve or may make it necessary to take also an epistemic and practical normative stance. This is what makes taking a normative stance such a complicated affair. As we noted earlier the "standards of excellence" of engineering practice cannot be dissected into independent sets of values and norms that correspond to the neatly defined values and norms of standard philosophical domains.

Third, if the philosophy of technology is taken to be itself a normative practice that is based on a substantive moral stance, then the question about its basic (morally relevant) ends and values arises. What are these values? For medical practices they are about curing diseases and improving the health of patients. For engineering practices these basic values are contained in their codes of conduct which usually contain phrases like "Engineers shall hold paramount the safety, health and welfare of the public in the performance of their professional duties".[12] In analogy to the ABET, that defines engineering as a profession for "the benefit of mankind", the philosophy of technology may conceive of itself as a discipline for the benefit of mankind. But, of course, this is not of much help; it does not tell us anything about how this basic value shapes the discipline.

If the core activity of a normative philosophy of technology is critical reflection on technology and taking a reflective or substantive normative stance on the issues is at stake, then the epistemic, practical, moral etc. values and norms underlying critical reflection itself have to be critically questioned at the meta-level. The result does not necessarily have to be a set of (morally relevant) substantive values and norms that have to be adopted by the discipline; these values and norms may also be of a procedural nature and involve a procedural role for philosophers of technology.[13]

A fourth and final point, related to the foregoing one, is that a normative axiological turn in the philosophy of technology will inevitably run up against what is known as the is-ought problem or as Hume's guillotine. This problem is the more pressing the more normative philosophical commitments are involved. In that sense it may affect the reflective position less than the substantive position. So far, we

---

[12] See, for instance, the code of conduct of the NSPE: http://www.nspe.org/resources/ethics/code-ethics, and of the ASME: https://community.asme.org/colorado_section/w/wiki/8080.code-of-ethics.aspx.

[13] Here we touch upon the ethics of philosophy as a profession; see (Hansson 2015).

have assumed that a descriptive axiological turn is a preliminary, necessary step for the normative axiological turn. The idea behind this assumption is that taking a well-informed normative stance to engineering practice – well-informed about the actual values and norms of engineering practice – is better than just taking a normative stance. In many normative practices, including engineering itself, the relevance of facts for taking a normative position or for making decisions that are normatively relevant goes undisputed. Nevertheless, there is the philosophical problem about how to argue from facts, independently of whether or not the content of those facts concerns values, to what ought to be done. For Hume this is logically impossible, because the conclusion contains a notion (the 'ought') that is not contained in the premises (the 'is') of the argument (Hume 1969 (1739–1740), p. 521). Hume's argument is not that facts may not be relevant for arriving at normative conclusions; in combination with normative premises they clearly are. His point is that from facts alone, no normative conclusions can be drawn.

So, why is Hume's guillotine a problem for a normative axiological turn in the philosophy of technology? Suppose that the discipline succeeds in spelling out its basic values and norms involved in its normative axiological turn (see the previous point). Then the assumption about the relevance and necessity of the descriptive axiological turn is not problematic; it may provide relevant facts for normative arguments. However, now the problem shifts to the grounding of the discipline's basic values and norms. In the axiological turn these values and norms are going to be the standards against which the standards of excellence of engineering practices are to be judged. How can this special status of the values and norms of a normative philosophy of technology be justified? If both a normative philosophy of technology and engineering have benefitting humankind as (one of) their ultimate end(s), what is it then that makes the values and norms of the normative philosophy of technology so special? Is there somehow a hierarchy of practices that leads to a hierarchy in standards of excellence?

When we take a closer look at how to justify the values and norms of a normative philosophy of technology, we see Hume's problem surfacing again. Schematically there are two approaches: they may be justified in an a priori way or by reference to factual circumstances. Both approaches raise their own problems. Apart from whether in general values and norms can be justified a priori, it is questionable whether a priori values and norms are useful as standards for judging the standards of excellence of engineering practices. It is to be expected that such a priori, timeless values and norms are very general in nature, that is, are *thin* ethical/normative principles (see Sect. 2.4). In order to be able to confront them with and to use them as standards for the more detailed, historically grown (and changing!) standards of excellence in engineering practices, these general values and norms will have to be specified and operationalized. In order to do so reference to factual circumstances will have to be made. References to these factual circumstances do not function merely as premises in an argument from which from general values and norms more specific ones are logically deduced. Much more appears to be involved in their role, since the specification and operationalization themselves involve normative decisions about whether a proposed specification and operationalization is morally acceptable in those particular factual circumstances (Kroes and van de Poel 2015). If that is indeed the case, then reference to those factual circumstances appears to justify the moral acceptability of the specification and operationalization of the values involved in which case we run again into Hume's guillotine. The second approach, in which the values and norms of a normative philosophy of technology are justified on the basis of factual circumstances will face the is-ought problem directly (leaving aside the possibility of moral realism).

Various proposals have been made to avoid the above dilemma. One of the most well-known and influential is Rawls' method of reflective equilibrium that combines (a priori accepted) general moral principles with factual, considered moral judgments (Rawls 1971). Less well-known is Dewey's attempt to ground the 'ought' in the 'is' (1891, p. 198): "I should say (first) that the "ought" always rises from and falls back into the "is," and (secondly) that the "ought" is itself an "is," – the "is" of action." For Dewey the separation of the "ought" from the "is" has had the catastrophic consequence of stiffening it "into a rigid external must, imposed no one knows why or how" (1891, p. 201). The problem we are facing here is again the one of how to relate in a coherent way moral (normative) judgments with factual ones. For the moment a notion of coherence that is able to adequately deal with this problem is still missing.

All in all, the foregoing points illustrate that, whereas a descriptive axiological turn is nothing more than an amendment to the empirical turn, a normative axiological turn in the philosophy of technology is of a much more radical nature in that it involves giving up the 'value-free' approach underlying the empirical turn.

2.6 Discussion: What to Aim For?

We have argued that in order to understand engineering practice as value-laden practice in all its richness the empirical turn has to be followed by what we have called a descriptive axiological turn. We have also pointed out that it may be difficult to implement a purely descriptive axiological turn. Furthermore we have analyzed what it means for the philosophy of technology to take a normative axiological turn. Up till now, however, we have been rather non-committal about whether it is desirable to take a normative axiological turn. Even if it turns out to be difficult to implement a purely descriptive axiological turn it may be taken as a 'value-free' ideal to strive for in the philosophy of technology, instead of embracing a normative axiological turn. In this final section we will address the issue whether the philosophy of technology should aim for a normative axiological turn or not. In MacIntyre's terms this is an issue about what are the defining features of philosophy of technology as a practice: what are its internal goods and it standards and norms of excellence.

If we conceive of the internal good of the philosophy of technology as in some way contributing to a better development/implementation of technology in our society, then aiming for a purely descriptive axiological turn appears problematic. In that case the philosophy of technology is vulnerable to the same charge that Winner (1993) has leveled against social constructivist theories of technology in his article *Upon Opening the Black Box and Finding it Empty*: *Social Constructivism and Philosophy of Technology*. Indeed, a descriptive axiological turn leaves the philosophy of technology without "anything resembling an evaluative stance or any particular moral or political principles that might help people judge the possibilities that technologies present" (Winner 1993, p. 371). Although it may certainly contribute to a better understanding of the values and norms that shape modern technology, it will leave us empty-handed when it comes to questions of how to deal with normative issues (moral or otherwise) with regard to modern technology. To avoid Winner's charge a normative axiological turn appears in order. This raises two questions. First, is such a turn desirable and, second, is the philosophy of technology up to such a task?

Let us first focus on the question whether a normative axiological turn is desirable. Recently, various philosophers of technology have staked the claim that it is the task of the philosophy of technology to contribute to a better development and implementation of technology. By doing so they implicitly or explicitly are arguing for what we have called a normative axiological turn. We will have a brief look at two examples.

Peter-Paul Verbeek (2010) has argued that the philosophy of technology should head for what he calls 'accompanying technology'.[14] According to Verbeek we have witnessed in the first decade of the twenty-first century an outburst of ethical studies of technology to which he refers as the 'ethical turn' in the philosophy of technology. In his opinion, however, most of these studies do not take into account the results of the empirical turn. He argues that now it is time for a third turn which has to integrate the empirical and ethical turn in order to develop an ethics of accompanying technology. The latter has as its central aim "to accompany the development, use and social embedding of technology" (Verbeek 2010, p. 52). It is an activist form of philosophy of technology in the sense that the philosopher of technology has to immerge herself in engineering practice (*ibidem*):

> Accompanying technology development requires a thorough engagement with designers, looking for points of application for moral reflection and anticipating the social impact of technologies-in-design. Rather than placing itself *outside* the realm of technology, this type of ethics explicitly engages with its development.

Instead of being a bystander who analyses the moral dimension of technology, the philosopher of technology should change her role to *doing* ethics of technology (p. 51).

In our terminology, Verbeek argues for a substantive normative axiological turn; the philosopher of technology becomes a direct participant in the development of technology. There is, however, a difference in scope. Verbeek's normative axiological turn only concerns ethical norms and values, whereas in our opinion such a turn has to involve other kinds of norms and values as well, in particular methodological

---

[14] He made the same plea in his presidential address to the 2015 SPT meeting in Shenyang (China).

and epistemological ones. This difference in scope is significant when thinking about the nature of the axiological turn and its implications. It is the difference between an *ethicist* and a *philosopher* becoming involved in technology development.

Also Philip Brey (2010) has addressed the question where to go with the philosophy of technology after the empirical turn.[15] He argues, again in our terminology, not only for a descriptive axiological turn, but also for a normative one. On the one hand he is in favor of a modest normative turn, one which comes close to what we have termed the reflective version of the normative axiological turn. In his opinion one of the three main questions of the philosophy of technology is: "How can the consequences of technology for society and the human condition be understood and evaluated?" (p. 43). On the other hand he also keeps open the possibility for a more or less direct involvement of the philosopher of technology in technological development, that is, for a substantive version of the normative axiological turn.[16] For him the philosophy of technology also has to address the following question: "How ought we to act in relation to technology?" (p. 43). This is what he calls the field of technology ethics. What is needed in this field are good theories about the moral agency of technical artifacts and ethical theories of human agency mediated through technology. On top of that, theories and methods of ethical technology assessment will have to be developed in order for "ethicists to have a constructive role in the assessment and *development* of new and emerging technologies" (p. 47; emphasis ours). Finally we need in this field "better methods for the ethical analysis and *guidance* of social and political debates surrounding the introduction of new technology" (*ibidem*; emphasis is ours). So, ethicists will be directly involved in and influence on the one hand the development of new technology but also, on the other hand, public debates about new technologies. This is very much in line with our substantive version of the normative axiological turn.

Apart from Verbeek's and Brey's calls for an ethical turn, we are witnessing many more initiatives in the field of the philosophy of technology that aim at a more direct involvement of philosophers of technology in technology development and implementation. Suffice it here to mention the center for Value Sensitive Design (http://www.vsdesign.org/), the Center for Science and Technology Policy Research (http://sciencepolicy.colorado.edu/about_us/find_us.html),[17] the Design for Values initiative (van den Hoven, Vermaas, and van de Poel 2015) and the 3TU. Center for Ethics and Technology (http://ethicsandtechnology.eu/). What these initiatives have in common is that they seek close contact with and to some extent immerse themselves in the phenomena and practices they study.

Most of these initiatives involve taking a normative stance and therefore implicitly or explicitly endorse a normative axiological turn. We very much share with these initiatives the underlying motivation that the philosophy of technology should

---

[15] Brey distinguishes between two different empirical turns; this distinction, however, is not relevant for the following, so we will ignore it here.

[16] See also his contribution to this volume.

[17] See also Briggle's call for a policy turn in the philosophy of technology in this volume.

strive to contribute to societal issues and to a technology that helps to solve them. However, anyone arguing for a normative axiological turn in the philosophy of technology will have to face the problem of the justification of its normative stance. It will not do to simply assume that a philosophy of technology that implements a normative axiological turn will improve the development of technology and the societal debates about technology. In our opinion a normative axiological turn will have to be a self-reflective process in which the values and norms of the philosophy of technology, not only as an academic activity, but also as a practice involved in technology development, will have to be object of critical reflection. Neither Verbeek, nor Brey explicitly address this issue in their call for an ethical turn. In our opinion it is a necessary step and a real challenge that anyone advocating a normative axiological turn for the philosophy of technology will have to face.

Apart from justification of its normative stance we have to ask ourselves on what grounds (knowledge, expertise, skills) the philosophy of technology may claim to be able to contribute to better technologies in our society. As we remarked above, a normative axiological turn will have to build on a descriptive axiological turn. For making evaluative judgments about what is going on in an engineering practice, let alone normatively interfering in that practice, knowledge of the relevant facts and understanding of that practice is necessary. But how detailed and how reliable should this knowledge and understanding be before one is justified in taking a normative stance? Is it indeed the case that the philosophy of technology has reached a level of maturity that it may justifiably and confidently take a normative stance and make good on its claims or promises? On what specific philosophical knowledge and skills are these claims and promises based?

These are thorny questions. They touch upon whether the philosophy of technology is a form of theoretical philosophy or of practical philosophy, or both; and if both, how these two forms are related to each other. It not only involves issues about internal and external goods of the philosophy of technology, but also about its standards and norms of excellence. In short it involves what the philosophy of technology is all about, that is, its identity is at stake. As a theoretic discipline the internal goods of the philosophy of technology and its standards of excellence are mainly related to understanding technology and its role in modern life. As a practical discipline its internal goods and standards of excellence are related to contributing to 'better technology'. Of course, the internal goods of a theoretical philosophy of technology (understanding technology) may contribute to bringing about better technology but in so far this is the case it is an external good produced by that practice.[18] So, what is the internal good of a practical philosophy of technology is an external good for its theoretical counterpart. Given this difference in internal goods it is not to be expected that the same standards of excellence are operative in both

---

[18] Note that our notion of an external good differs from MacIntyre's notion; according to MacIntyre (1984, p. 188) external goods are "contingently attached" to practices and can be achieved in alternative ways. This is not the case for external goods that are realized by putting an understanding of technology to practical use in bringing about better technology.

fields.[19] If not, then it may be questioned whether we are dealing with one and the same practice.

A brief comparison with the difference between physics and engineering physics may be illuminating. The internal goods of physics may be described as knowledge and understanding of the physical world and through the peer review system contributions to this good are judged. The task of the peers is to guard the standards of excellence of physics. The internal goods of engineering physics may be described roughly as useful knowledge for the design of technical artifacts and designs of new technical artifacts. It has its own peer review system for judging contributions to this internal good. Although there may be some overlap between the two peer review systems, it is clear that the standards of excellence in both fields are not the same. Nevertheless, history shows that there may be a strong interaction between both fields and that they can benefit a lot from each other.[20]

The situation with regard to theoretical and practical philosophy of technology may be similar. They may be considered distinct philosophical practices that nevertheless benefit from each other. This way of looking at things underscores that there is a radical difference between a descriptive and a normative axiological turn in the philosophy of technology. We have already pointed out that the change from the position we advocated in the empirical turn to a descriptive axiological turn is an innocent one in the sense that it implies only a further specification of the object of study. This does not involve a change in internal goods nor in the standards of excellence. In so far mainstream philosophy succeeds in spelling these out clearly, the philosophy of technology can take its bearings with regard to these from mainstream philosophy. With regard to a normative axiological turn in the philosophy of technology still a lot of work remains to be done in setting and explicating its standards of excellence; in this respect the philosophy of technology may learn from other fields in applied philosophy.

In summary, we think that the philosophy of technology should take the issue of the relevance of the internal goods its delivers seriously, not by rushing headlong into a normative axiological turn, but by addressing the challenges that such a turn poses to the field. If indeed one of its defining internal goods is taken to be to contribute to better technologies in our society, then a call for normative axiological turn is in order. In answering to this call it will have to clarify and justify its own epistemic, practical and moral values and norms and its own standards of excellence and how they relate to the values and standards of excellence of other stakeholders involved in realizing better technologies. In other words, it will have to redefine itself as a philosophical endeavor and it has to answer the question what kind of

---

[19] MacIntyre (1984, pp. 188–189) writes that internal goods "can only be identified and recognized by the experience of participating in the practice in question. Those who lack the relevant experience are incompetent thereby as judges of internal goods." Thus the standards of excellence for judging the realization of internal goods appear to be closely related to the nature of those internal goods.

[20] Houkes' analysis of the interaction of basic and applied research on Perovskite solar cells may be interpreted as a nice example of the subtle interplay between internal and external goods in the development of physics and engineering; see Houkes' contribution to this volume.

stakeholder it wants to be in the development, design and production of new technologies.

References

Borgmann, A. (1995). Does philosophy matter? *Technology in Society, 7*(3), 295–309.

Brey, P. (2010). Philosophy of technology after the empirical turn. *Techné: Research in Philosophy and Technology, 14*(1), 36–48.

Dewey, J. (1891). Moral theory and practice. *International Journal of Ethics, 1*, 186–203.

Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science, 67*, 559–579.

Hansson, S. O. (2015). The ethics of doing philosophy. *Theoria, 81*, 93–96.

Hume, D. (1969 (1739–40)). *A treatise of human nature*. Harmondsworth: Penguin.

Jeffrey, R. C. (1956). Valuation and acceptance of scientific hypotheses. *Philosophy of Science, 23*, 237–246.

Kirchin, S. (Ed.). (2013). *Thick concepts*. Oxford: Oxford University Press.

Kroes, P. (2006). Coherence of structural and functional descriptions of technical artefacts. *Studies in History and Philosophy of Science, 37*(1), 137–151.

Kroes, P. (2012). *Technical artefacts: Creations of mind and matter*. Dordrecht: Springer.

Kroes, P., & Meijers, A. (2000). Introduction: A discipline in search of its identity. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (pp. xvii–xxxv). Amsterdam: Elsevier Science Ltd.

Kroes, P., & Meijers, A. (Eds.) (2006). Special issue on the dual nature of technical artefacts. *Studies in History and Philosophy of Science*, Vol 37

Kroes, P., & van de Poel, I. (2015). Design for values and the definition, specification, and operationalization of values. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of ethics, values, and technological design* (pp. 151–178). Dordrecht: Springer.

Kroes, P., & Verbeek, P.-P. (Eds.). (2014). *The moral status of technical artefacts*. Dordrecht: Springer.

MacIntyre, A. C. (1984). *After virtue: A study in moral theory* (2nd ed.). Notre Dame: University of Notre Dame Press.

Meijers, A. (Ed.). (2009). *Philosophy of technology and engineering sciences*. Amsterdam: Elsevier.

Mitcham, C. (2014). The true grand challenge for engineering: Self-knowledge. *Issues in Science and Technology, 31*(1), 19–22.

Quine, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review, 60*(1), 20–43.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rudner, R. (1953). The scientist *qua* scientist makes value judgments. *Philosophy of Science, 20*, 1–6.

van den Hoven, J., Vermaas, P. E., & van de Poel, I. (Eds.). (2015). *Handbook of ethics, values, and technological design; sources, theory, values and application domains*. Dordrecht: Springer.

Verbeek, P.-P. (2010). Accompanying technology: Philosophy of technology after the empirical turn. *Techné: Research in Philosophy and Technology, 14*(1), 49–54.

Williams, B. (1985). *Ethics and the limits of philosophy*. London: Fontana.

Winner, L. (1993). Upon opening the black box and finding it empty: Social constructivism and philosophy of technology. *Science, Technology & Human Values, 18*(3), 362–378.

Peter Kroes Engineering degree in physical engineering from the University of Technology Eindhoven and Ph.D. in the philosophy of physics from University of Nijmegen (The Netherlands). He is professor in philosophy of technology at the University of Technology Delft. His main areas of interest are philosophy of technology and philosophy of science. Recent book publications: *Artefact kinds; ontology and the human-made world*, eds. Maarten Franssen, Peter Kroes, Thomas A.C. Reydon, Pieter Vermaas, Synthese Library 365, Springer, 2013, *Technical artefacts: creations of mind and matter*, Springer, 2012.

Anthonie Meijers Studied mechanical engineering at Delft University of Technology and philosophy at Utrecht University. He has a Ph.D. in philosophy from Leiden University. He is full professor at the Department of Philosophy and Ethics of Eindhoven University of Technology. His main interests in the field of the philosophy of technology are the theory of artefacts, agency and artefacts, technology and behavioral change, and the epistemology of technology. He is one of the co-editors of the 2000 book on the empirical turn in the philosophy of technology and the editor in chief of the *Handbook Philosophy of Technology and Engineering Sciences* (2009). He is also editor and co-founder of Philosophical Explorations, an international journal for the philosophy of mind and action.

Chapter 3
Philosophy of Technology as a Serious Branch of Philosophy: The Empirical Turn as a Starting Point

Maarten Franssen and Stefan Koller

3.1 Introduction

This contribution departs from the observation that, notwithstanding the empirical turn, work in the philosophy of technology is still too fragmented and isolated, both internally, in how its various themes are mutually related, and externally, in how well its themes are linked up to what happens in the established major fields that make up philosophy as a whole. We argue that the philosophy of technology as currently practiced has to extend both in scope and method and that a systematic exploration of its connections with the core fields of philosophy will help it develop into a mature field. We, then, mount a diagnosis and provide a remedy. Greater systematicity is *needed* to counteract the fragmentation and lack of substantive unity in philosophy of technology. Such systematicity can be *provided* by means discussed in this paper, namely, by checking the content and validity of new contributions against both extant results in philosophy of technology and (conceptually or inferentially) related positions in foundational analytic philosophy, above all metaphysics, epistemology, and the philosophy of language.

Our paper eschews detailed argument for the diagnostic part. We take the observation of such lack of unity and systematicity to be a relatively uncontroversial empirical claim, confirmed by personal experience and peer testimony.[1] Insofar as

---

[1] Arguably the *content* of the observation is controversial – what it is held to imply about the state of the field and what can and should be done about it – but not its *truth value*. Our observation's authority further rests on the field's entry in the *Stanford Encyclopaedia of Philosophy*, co-curated by one of us since 2009 Franssen et al. 2013). That entry presents the field *in its present state* in as systematized a manner as possible, with the limitations of that systematization clearly in view: an *inventory* of issues and positions does not make a field, but at most present the making of one.

M. Franssen • S. Koller (✉)
Department of Philosophy, Delft University of Technology, Delft, The Netherlands
e-mail: m.p.m.franssen@tudelft.nl; s.koller@tudelft.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_3

we argue for it, it is through pointing out, in the remedy part, what potential there is for a more developed philosophy of technology that currently goes unnoticed and unpractised. But nor do we address arguments that contest the suitability and success of our proposed remedy, since such arguments do not presently exist and are, at any rate, better explored and articulated once their target has been properly developed to permit precise objection and refutation.[2] It is our paper's primary objective to furnish readers with exactly that target, and to illustrate by concrete examples the type of remedy – the type of doing philosophy – we claim philosophy of technology requires.

Our first two sections provide the diagnostic part. We chart major points of debate, and outline how its fragmentation and isolation have configured the field so far, and especially its relation to philosophy of science. Sections 3.4 and 3.5 explore its (potential) connections to other areas of philosophy, organized around a basic conception of the core themes of philosophy of technology, here proposed as three themes: (1) the nature of artefacts; (2) the concept of design; (3) the notion of use. Section 3.6 demonstrates with respect to (1) and (3) that stronger ties to contemporary research in metaphysics, epistemology, the philosophy of language, and (meta-) ethics, are necessary and rewarding – thus demonstrating our prospects toward a 'unified philosophy of technology', one constrained by the meta-philosophical virtues outlined in our essay's earlier sections.

3.2 At the Empirical Turn

In their editorial introduction to the volume *The empirical turn in the philosophy of technology*, Peter Kroes and Anthonie Meijers described the philosophy of technology as a discipline in search of its identity. They let follow this observation by a characterization of philosophy of technology as practiced from the 1960s to the 1990s, as "dominated mainly by metaphysical analyses of technology (under the influence of Heidegger), and by critical reflections on the consequences of science and technology for the individual and social form of life" (Kroes and Meijers 2000b, p. xvii). The identity crisis should perhaps rather be interpreted as a dissatisfaction of many scholars philosophically interested in technology (including the volume's editors) with the approach to philosophy of technology dominant until the 1990s. The common complaint was that philosophy of technology did not take technology itself – more precisely the activities of engineers and the conceptual and tangible

---

[2] That is, even readers agreeing with our diagnosis may wish to explore means of remedy other than the ones we provide here. We would be the first to welcome the ensuing methodological diversity – a diversity disciplined by a shared metric of success, namely that of systematicity. As Williamson (2007, pp. 285–286) remarks a propos *disciplined* methodological diversity in philosophy, "Tightly constrained work has the merit that even those who reject the constraints can agree that it demonstrates their consequences."

output of these activities – sufficiently seriously. Or to put it more starkly, the complaint was that philosophy of technology entirely ignored this dimension of technology, with detrimental effects to the quality of its analyses as contributions to philosophy. Philosophy of technology could not be taken seriously as a philosophical discipline as long as it did not notice and take into account the engineering dimension of technology, which is responsible for the coming-into-existence of technology and has a large, if not decisive, say in what it comes to look like. A reorientation of the field was necessary, of the form of an "empirical turn".

In their preface to the same volume, the editors expressed high hopes for what this reorientation of philosophy would deliver (Kroes and Meijers 2000a):

> Philosophers have to open the black box of technology and base their analyses on an inside view of the engineering practice and on empirically adequate descriptions of technology. Such an empirical turn will not simply provide illustrations of existing philosophical ideas and analyses, but will open up whole new areas of research. Methodological, epistemological, ontological, and ethical questions concerning technology and the technical sciences will eventually come into focus that so far have been addressed only marginally or not at all in philosophy of technology.

It is this "coming into focus" of methodological, epistemological, ontological and ethical questions concerning technology that we wish to look into more closely in this essay. Kroes and Meijers started their editorial introduction with a characterization of philosophy of technology as professed in the 1960s–1990s, but exactly as a field of philosophy its position was an extremely marginal one. Much of it, notably the "critical reflections on the consequences of science and technology for the individual and social form of life", was probably not seen as being serious philosophy – where 'serious philosophy' is understood as philosophy practised as an academic profession – or as belonging to philosophy at all. What is more, earlier attempts to delineate a philosophy of technology that took engineering seriously, which there definitely were prior to the 1990s, did not occur within philosophy, including philosophy of science. The ground-breaking exchange of ideas between Bunge, Mumford, Skolimowski, Jarvie and others on how to differentiate technology from science appeared in the journal *Technology and Culture* (Kranzberg 1966), which was dedicated primarily to historical studies related to technology and was never a philosophy journal.

3.3 Since the Empirical Turn

Have the high hopes for the philosophical maturation of philosophy of technology materialized? An answer to this question has a substantial and an institutional dimension. With regards to the latter, there is some progress, but perhaps not altogether in the direction envisaged in 2000. Since 2011 Springer publishes a journal *Philosophy and Technology*, but it may be thought significant that its title is not *Philosophy of Technology*. Due to the many intimate relations and similarities between the practices of science and engineering, philosophical analyses of technology that take engineering seriously naturally take philosophy of science, with its impressive pedigree within philosophy, as a benchmark of sorts.

Indeed, one of the methodological questions that the empirical turn has generated is whether (*a*) a mature philosophy of technology has the philosophy of engineering as one of its constituents, or even its major constituent, or whether (*b*) the philosophy of engineering should be considered a new or emerging field, which has some overlap with the philosophy of science on the one side and has different overlap with (various types of) the philosophy of technology on the other side. Some institutional developments have taken place recently that favour the first option: the desired reorientation of philosophy of technology to bring engineering into view occasionally seems to take the form of a gradual reconstruction of philosophy of technology as philosophy of engineering, under the general heading 'philosophy of engineering and technology' which is as unclear as it is uncommitted. In 2010, Springer also started a book series called *Philosophy of Engineering and Technology*, under the chief-editorship of Pieter Vermaas, in which the present volume is published as its 23rd title. Further, one of the editors of the "Empirical turn" volume subsequently acted as editor of the volume entitled *Philosophy of technology and engineering sciences*, issued as part of the 16-volume *Handbook of the philosophy of science*, published by North-Holland between 2006 and 2012 (Meijers 2009). And between 2011 and 2015 the Division of Logic, Methodology and Philosophy of Science of the International Union of History and Philosophy of Science, organizer of the esteemed International Congresses of Logic, Methodology and Philosophy of Science since 1960, changed its name to Division of Logic, Methodology and Philosophy of Science and Technology, to open up its congresses as a forum for work in the philosophy of technology and of engineering, in response to pressure from the side of 'empirically-turned' philosophers of technology.

This may serve as a relatively cheap entrance ticket to the continent of respected philosophy, but it is also a relatively modest one, when we review the high hopes expressed in the preface of the *Empirical turn* volume. The field in which the philosophy of science holds central position has always favoured methodological and epistemological questions over ontological and ethical ones, and to choose the philosophy of science as a benchmark may constrain the philosophy of technology in developing its full potential. The undeniably intimate connections between the practices of science and engineering should not block from view that there are major differences in orientation between science and technology. Arguably one of the most blatant differences is that science does not aim to influence how people live their lives, whereas engineering strenuously aims to do so. Ethical and, plausibly, also ontological questions seem to have their relevance precisely in relation to this point where science and technology part directions.

With these considerations we have already started addressing the substantial dimension in answering our opening question. The past fifteen years have seen contributions that address methodological, epistemological, ontological and ethical questions which were either (*a*) raised in the course of investigating specific features of or problems of engineering design or (*b*) raised with designing engineers as their main targets, or that were (*c*) addressed with an emphasis on technology and artefacts as intentional products of engineering design. Although this work undeniably has contributed to a reorientation of work done under the heading of philosophy of technology, it is questionable whether all together it amounts to a satisfactorily reoriented philosophy of technology. Bringing the engineering dimension of technology into view is not sufficient for elevating the various philosophical analyses of technology to a philosophical discipline of a maturity and standing comparable to, for example, the philosophy of science or, for that matter, the philosophies of several other practices, such as the philosophy of art and the philosophy of language. If philosophers generally are to stop looking upon the philosophy of technology as, at most, a subfield of the philosophy of science, and instead should come to accept it as a field in its own right, and the practice to which it is dedicated as worthy of the corresponding level of philosophical analysis, then the philosophy of technology will have to develop beyond the work induced by the empirical turn.

Finally, the 'substantial' and 'institutional' factors we mention tend to be mutually reinforcing – which would explain both the factors' hold on the field and the field's subsequent inertia. That is, philosophy of technology's lack of substantive unification is (at least in part) a consequence of its institutional fragmentation, *and* vice versa. Philosophy of technology is pursued in geographically and culturally extremely diverse types of philosophy. To high-light this (laudably) ecumenical outlook seems to have been a major editorial concern in another recent 'taking stock', that of Blackwell's *Companion to the philosophy of technology* (Berg Olsen et al. 2009). The concern's immediate result is to accommodate as many 'types' of philosophy as possible, without the slightest attempt to have these types talk to (let alone, constrain) one another. That is, the field presented in that *Companion* is divided, not simply by a list of issues or positions, but by types of philosophy one can enlist in studying (the various guises of) technology. Thus, 'phenomenology' merits a distinct entry, 'analytic philosophy' another, and so on.

The end result not only departs from the editorial practice of the volumes published under 'the empirical turn', but from editorial practice in 'Philosophy of…' companions more widely. Where other 'philosophies of…' may at times give the impression of a field in search of its subject, contemporary philosophy of technology (as summarized in Blackwell's *Companion* volume) rather appears to be a subject in search of a reasonably unified philosophy. This is where the present paper comes in. And again, our point is that factors that presently hold back philosophy of technology have been overcome by mature 'philosophies of…'. In philosophy of art, for instance, philosophers of just as diverse orientations engage a shared set of topics, a shared set of key figures (like Kant and Hegel), and a shared set of 'examples' (canonical artworks). This demonstrates that a field's (warranted degree of) being 'ecumenical' and its being (reasonably) unified – with results generated under mutual exchange and tested for coherence to previously reached results – need not be mutually exclusive.

3.4 Beyond the Empirical Turn

If our diagnosis is even remotely correct, philosophy of technology as currently practiced has to extend both in scope and method if the hopes expressed with the presentation of the empirical turn are to be realized. Only a systematic exploration of what ties philosophy of technology has with the core fields of philosophy will help it develop into a mature field – with the plausible outcome that, once it has gained its place as a mature field, these ties will settle themselves, some as strong and others as weak, shaped by (*i*) the largely internal research questions that characterize mature practice-oriented fields like the philosophy of science or the philosophy of art as well as the core philosophical fields like metaphysics or (meta) ethics, and by (*ii*) the sort of awareness that maturity enables of how these research questions and the work they generate are relevant to each other.

To explore these ties, a basic conception of the core themes of philosophy of technology is necessary. We begin by providing such a thematic conception or framework ourselves (Sect. 3.4.1), then compare how our framework fares in relation to recent competitors (Sect. 3.4.2), and address a major objection our framework has to face (Sect. 3.4.3). Section 3.5 then explores ways in which each of our three core themes can be connected to main areas in philosophy that we consider promising and that have indeed already been partially investigated. Section 3.6, by far the longest, supplements our threefold framework by a ‘use-theoretic’ proposal that will demonstrate the paper’s positive goal to provide concrete steps towards a more systematic philosophy of technology.

3.4.1 Thematic Framework

We propose as forming the thematic core the following three themes:

1. The nature of artefacts;
2. The concept of design;
3. The notion of use.

We do not claim that these themes jointly exhaust what distinguishes philosophical reflection on technology from reflection on other practices, but we think that they are each broad enough and basic enough to jointly cover most of what is relevant to the field. In this way they serve to roughly characterize the field and to delineate it sufficiently from the philosophical study of neighbouring practices, such as science and art. And of course, some topics *seemingly* excluded by (1) to (3) can be brought into their fold later on, by way of derivation. One notion in particular that could be considered fundamental enough to deserve inclusion in our list of themes is ‘making’. However, we take it that technology is concerned with a specific form of making – plan and design-based – only, and that form can and will receive consideration by indirect means. For instance, our Sect. 3.6 brings together (1) and (3) in a sketch of a metaphysics of artefacts based on the 'use-relevant properties' of artefacts. From there, a next step to be explored might be to define the *making* of artefacts in terms of a certain practical knowledge: knowledge enabling the maker (engineer) to *see to it that* the respective artefact ends up having the 'use-relevant properties' that define its type (see Sect. 3.6.3 for more on this).

The choice of these three themes for demarcating the field of philosophy of technology amounts to the philosophical claim that each of the three concepts of artefact, design and use is central to technology in a way that it is not – in particular not when that concept's relations to the other two concepts are taken into account – central to any other practice. This may be easily granted for 'use', if only for the reason that the notion is so nearly absent from the philosophical literature that no protest is likely to be raised. For 'artefact' and 'design', however, our claim of exclusiveness seems questionable: should not both, for example, be granted a central position in the philosophy of art?

This can be quickly denied for 'design'. Although all artworks may be conceived as coming about through a creative process of a clearly mental nature, the philosophy of art seems not much interested in it. Just as is the case for hypotheses and theories according to the logical-empiricist view of science, the process of the generation or creation of the artwork belongs to the context of discovery, which is the context in which the artist or scientist can play around still free from philosophical scrutiny. Philosophy becomes interested once the period of gestation has resulted in an object of interest. To be sure, in the philosophy of science this point of view is now an old-fashioned one: the heuristics of hypothesis and theory formation is certainly a topic for study now, because it was never plausible that the research phase was open to that much free play, even if science could live with occasional whimsicality. This is even more true for technology and engineering, given the very real constraints of time and money there, and accordingly the process of hypothesis and design-concept formation is seen as under severe pressure of answering to a substantial minimum of rules, method and reason. The notion of 'design' refers to such a process. Art seems exactly to be that practice that is not subject to the pressure to organize its creative activity in the form of design, that is, answering to rules, method and reason. Even if artists of all stripes have sought the support of 'ideas' for their work, such support typically is individual to the extreme. Research on the ontological status and proper delineation of 'artwork(s)' will occasionally (say, in contrasting institutionalist to non-institutionalist delineations) *mention* the progeny of lone creative will, but not much philosophical attention is assigned to the specifics *behind* such progeny.[3]

---

[3] It should be noted that it is an issue of considerable philosophical interest whether 'design' is necessarily an intentional or mental activity. At least one account has been proposed that construes the notion as broader than that, in the same way that there is a broad notion of function underlying both biological and artefact functions. In fact, the broad concept of design was construed exactly to ground the broad concept of function (see Krohs 2009). The further discussion of this issue does not fall within the scope of this paper, however.

With respect to ‘artefact’, indeed the few philosophical accounts that propose definitions of ‘artefact’ – the ones of Dipert (1993) and Hilpinen (1992, 1993) – propose broad definitions that include technical artefacts as well as artworks. Undeniably, our proposal creates in this way an overlap between the philosophy of technology and the philosophy of art, but we see no problem in this, as the presence of the two other themes, design and use, serve to distinguish technology sufficiently from art and neither of them seems relevant for understanding artworks. The overlap may even be considered a revealing one. Interestingly, neither art nor technology could be said to have the strongest claim on the notion of an artefact. Artworks are not necessarily objects – for example, theatre and musical performances are not – but neither are feats of engineering necessarily objects – as shown by chemical-process technology, software engineering, and systems engineering. For both art and technology it could be said that they are not fundamentally about making things but about making things happen, albeit different kinds of things. However, both rely for this on manufactured objects – tools, devices, instruments, props, costumes – to such an extent that the conceived and manufactured object is as good as constitutive of the practice (see further Sect. 3.6).[4]

3.4.2 Competing Frameworks

In remaining with artefact, design and use as the ‘defining terms’ of a mature philosophy of technology, a few words are in order on how this is different from other schemata to be found in the context of the empirical turn. Central to the research programme ‘The dual nature of technical artefacts’, led by Peter Kroes and Anthonie Meijers, is a scheme that singles out ‘(technical) artefact’, ‘function’, ‘(physical) structure’ and ‘human action’ as key concepts (e.g. Kroes 2012, p. 41). This is done, however, in the context of a programme that explicitly focused (for very good reasons, to be sure) on just one of our three themes, the nature of artefacts. Of the other two themes, the notion of use belongs to the sphere of human action. When developing an analysis of the nature of artefacts one indeed should not restrict *a priori* the spectrum of human actions in which artefacts may figure. However, as a key theme for structuring philosophy of technology generally, only a narrower conception of action, that of using things, makes sense. The second theme, in its turn, the concept of design, is linked to ‘function’, but straightforwardly corresponds to it just as little as ‘human action’ corresponds to ‘use’. Design is what underlies the fact that certain physical structures have *technical* functions. The concept of function as such, however, although not as broad as ‘human action’, has a much wider application, both

---

[4] The performative arts are the notable exceptions to this claim. On the complications such art forms create for a general ontology of art, and on how to overcome these, see Davies (2004). Whether there are similar exceptions on the technology side is much less clear, and insofar as there is indeed a difference, this may further serve to demarcate art from technology and, arguably, technology from ‘social engineering’, that is, all forms of social and societal interventions.

in relation to the natural world (biological functions) and to the social world (latent and manifest functions), and it is therefore key (partly in the form of ‘functionalism’) to several other philosophical fields, such as philosophy of science and philosophy of mind, which is one of the reasons why we do not propose it as a key notion for a philosophy of technology (in Sect. 3.5.3 a further reason is given).[5]

In his introduction to *Philosophy of technology and engineering sciences*, Meijers (2009) refers to yet another conceptual framework: the fourfold distinction of philosophical inquiries into technology proposed by Carl Mitcham in his *Thinking through technology* (1994). Mitcham distinguishes between conceiving of technology as object, as knowledge, as activity and as volition. Our approach as well as that of the editors of *The empirical turn* coincides (in substance if not letter) with Mitcham’s first three conceptions, but ‘technology as volition’ seems to have no place in it, just as it is omitted by Meijers when elaborating each of Mitcham’s headings. This narrowing down of technology compared to how its scope can be seen, and is seen by at least one prominent philosopher of technology, raises grave (meta) philosophical questions of its own, which we cannot properly pursue here. Suffice it to say that in excluding the arts as a ‘field’ or ‘practice’ which borders on technology and partly overlaps with it (witness the practice of architecture as a hybrid of these two), (post)empirical philosophy of technology severs links to the arts and humanities, and raises serious obstacles to the study of technology from those angles and subjects. Concerning our own approach, we show in Sect. 3.6 how construing artefacts ‘as expression of volition’ could be drawn in after all – and thus help mature recent attempts to (re)connect philosophy of technology to (core issues traditionally studied in) philosophical aesthetics.[6]

As stated, the purpose of the selection of our three key themes is not to propose and to defend them as a ‘definition’ of technology, as the outcome of a philosophical exercise in conceptual analysis aimed at revealing to us the ‘essence’ of technology. The purpose is rather to hook philosophical reflection of technology to as many fields of philosophy as seems relevant, and thereby both develop philosophy of technology into a coherent discipline and integrate this discipline within the overall field of academic philosophy. In doing so, we take the practice of technology (largely) for granted. In the next section we briefly discuss how justified this assumption is.

---

[5] The ‘Dual Nature’ research programme has resulted in a large literature on technical functions, including the so-called ICE theory of technical function (see esp. Houkes and Vermaas 2010). In our opinion, this literature has not established that the notion of function belongs primarily or first of all to technology. One might even argue that by developing a special theory for *technical* functions this work has done the opposite. Whether ‘function’ can count as a unitary concept and what unites its various uses remains an issue to be settled. To single out ‘technical function’ as a primitive concept structuring technology would amount to taking a position with respect to this issue, whereas one would rather hope that a mature philosophy of technology will contribute to clarifying it. As an aside, it could be remarked that the term came to technology and engineering later than to biology and social science, but this is not the place to document this claim.

[6] E.g. Coeckelbergh (2014, p. 46) and Illies and Ray (2016, pp. 83–85).

3.4.3 On (Not) Defining 'Technology'

Some readers may balk at our proposal's (Sect. 3.4.1) ease to take for granted the very concept that our philosophy is (purportedly) a philosophy *of*. Imagine that philosophy of religion or philosophy of science never explicitly attempted to demarcate their respective subjects (or themselves from each other). This raises the issue of what sort of 'demarcation point' will do.

Technology, science, art and religion are all ambiguous between (*a*) the practice strictly speaking, the activity of being engaged in doing science, technology, or the process of people being engaged in doing these things, and (*b*) the outcome or products of this practice. Under (*b*) we can then consider science as the sum total of scientific knowledge, technology as the sum total of designed, manufactured, operated and maintained devices, art as the sum total of created artworks, and religion as the sum total of deities and dogmas.[7]

Focusing on the product side (*b*), technology seems reasonably 'localized', not worse at least than science, religion, or art. Focusing on the activity side (*a*), however, technology seems a more problematic notion than any of the other three.[8] In our society, technology is organized overwhelmingly through privately owned, commercial firms that compete on many markets and whose prime driver is not the production of artefacts but the making of profit. As a result, technology as a practice blends into much, if not all, of the social, economic and political organization of our society. Plausibly for this reason, technology is often approached as first of all a product, the totality of designed, manufactured, operated and maintained devices. Technology as a practice might make too large a phenomenon, so to speak, and indeed there is a tendency to reinterpret it as the practice of engineering instead.

If there exists, then, this real (and to be reckoned with) danger of casting the 'extensional' net too widely, in a manner of speaking, the opposite danger introduced by the 'instead' in the above clause is just as real. As already discussed in Sect. 3.3, engineering narrows the practice down much too radically; it leaves little or no room for the *use* of technology and we would argue that no philosophy of technology can afford to do that. A narrowing down of technology as a practice to engineering is likely to continue and strengthen a development, noted above, of philosophy of technology, or rather philosophy of engineering, into a sister discipline of the philosophy of science, and arguably doom it to linger in its shadow forever. Nevertheless, any philosophy of technology conceived more broadly should be able to somehow incorporate a philosophy of engineering.

---

[7] Language is a practice for which the ambiguity between the activity and its product is more difficult to pin down. In Sect. 3.6, we discuss important connections between the philosophy of technology and the philosophy of language.

[8] Involved in this claim is the notorious English term 'technology', which merges an object and its study into a single notion. Many other languages are careful to make the distinction. In French and German, for instance, there is, still (although one may wonder how long it can resist the dominance of English), a significant difference between 'technique' and 'technologie' and between 'Technik' and 'Technologie', respectively.

A way to navigate the Scylla and Charybdis of our twin dangers is to tentatively treat engineering as a *paradigm case* of the practice of technology and its product, in the manner that Aristotle postulated 'core uses' for certain concepts like health, justice, and goodness. What makes such uses 'core' is that other, derived or derivative, uses of the same concept are explained *by reference to it* (but not vice versa). A person's healthy constitution is invoked in explaining the healthiness of a diet or form of exercise (but not vice versa).[9] We propose, as a working hypothesis, to think of 'engineering' as a paradigm or 'core use' of technology. If this hypothesis can be sustained then, as long as philosophers succeed in clarifying with respect to engineering the main philosophical questions about technology (quite a tall order already), extensional worries about 'technology' in 'philosophy of technology' can be settled later. In one scenario, results gained by studying the 'paradigm case' hold important philosophical lessons for (and are to be carefully 'applied to') a wider range of ('derived') phenomena. In the contrary scenario, lessons gleaned from our 'paradigm case' will neatly delineate the subject in its entirety. In neither case have we committed an extensional error.

This leaves a final worry, namely to what extent engineering *can* be treated as a paradigm case. There is something to be said for engineering being a hybrid practice, part 'pure' technology as we conceive of it here and part science. This invites the response that this hybrid practice of technology-as-engineering is the real one (to be studied philosophically) and that the 'pure technology' part in that hybrid we are scaffolding is illusory, an artefact, if you will, of our approach. Again, even if such a view (and objection to our project) can be sustained, which presently is far from clear, it does not doom our vision of a mature philosophy of technology to futility. After all, even on a hybrid approach, its two halves require careful philosophical investigation – and our approach provides just that. Whether the approach furnishes only half of the missing piece or all of it is a question that can be legitimately postponed until the first results are in. In this regard, our response to the last objection structurally mirrors the response we gave to the previous one. Worries about extensional adequacy are real but should not put a stop to philosophical investigation.

Let us then accept, or continue to accept, the practice of technology as given. We propose to conceive of this practice, for the purpose of philosophical inquiry, as structured by three key concept: artefact, design and use. This can link philosophy of technology in several ways with fields in philosophy generally, and potentially in two directions. In one way, discussion concerning issues related to the key themes, and their interrelations, connect to more general discussions of general philosophical import, and profit from taking from these discussions whatever concepts, approaches and theories turn out to be helpful. In an obvious way discussions of the nature of artefacts can be connected in this way and to some extent have already been so connected. In a reverse way, insights achieved within a philosophy of technology that is securely anchored in philosophy can find their way to other areas of philosophy to be of help there, in the way, for example, that ideas on the reduction

---

[9] See Shields (1999).

of concepts and theories from the philosophy of science were subsequently used and adapted in the philosophy of mind. By having to answer to forays made in other areas of philosophy, philosophy of mind became, and has remained to this day, a systematic and vibrant subfield.

3.5 Exploring the Three Themes

Let us briefly discuss in a bit more detail how, concerning each of our three key themes, discussions in the philosophy of technology could connect with main areas in philosophy (semantics, metaphysics, epistemology, ethics) so that the elaboration of these themes is enriched and can attain a level of quality that has been reached elsewhere, through profiting from what is available there, and so that the corresponding philosophical fields themselves are enriched by the input they receive from the *terra nova* of technology. Our discussion here serves two purposes: to both explore, in a preliminary fashion, the fuller potential of ‘points of connection’ with main areas in philosophy, *and* to document (in a somewhat pedestrian fashion) the extent to which such ties are currently made, and could be made further. The first purpose is then pursued in greater length, and unencumbered by ‘present state of the art’ worries, in Sect. 3.6, where we propose a completely new framework to weave our three themes together.

3.5.1 The Nature of Artefacts

With respect to the nature of artefacts, there is reason for great expectations from the start. On the one hand, at the core of one of the main research programmes associated with the empirical turn is a thesis that seems blatantly metaphysical: that artefacts have a dual nature. Whether that thesis should indeed have been interpreted metaphysically has, however, been questioned (Mitcham 2002). In the context of the programme, only one publication addressed metaphysics (Houkes and Meijers 2006). Since then, several of the researchers involved in that programme have edited a volume that explores more in depth the metaphysical problems concerning artefacts (Franssen et al. 2014). This volume focuses not so much on the metaphysical status of artefacts as on that of artefact *kinds*. Indeed many of the contributions imply that, however one construes the ontology of artefact kinds, their classification must follow principles that are entirely different from those governing the classification of natural kinds.

Artefacts certainly seem to play a central role in some of the oldest ongoing debates in metaphysics, notably the ship of Theseus with respect to the identity conditions for entities that undergo change. On a closer look, however, the emerging problems exist for all objects that survive (to some extent) replacement of their parts (which includes biological organisms as well as all artefacts) or all objects that are necessarily made of some material but still so loosely that many changes to the material need not affect the object *qua* object and vice versa many changes to the object need not affect the composing matter *qua* matter. The most interesting problem here is arguably first of all to identify the problems that artefacts *in particular* pose. Artefacts have been recognized as problematic for two contrasting views on 'ordinary' non-simple objects. One view is the modern continuation of the Aristotelian view that identity is identity under a sortal and that with each sortal is associated an internal unifying principle that secures a thing's identity over time. Artefacts seem exactly to lack such a principle, so much so that within this tradition the term 'artefact' has even been proposed as defining anything that lacks clear identity conditions, which does not refer at all to technology and design (Wiggins 1980, p. 89). Another view sees identity residing first of all in material unity. For that view artefacts pose a problem due to their compositionality: they can be disassembled and reassembled seemingly without this affecting their status as identifiable objects. Within this view, it has been proposed that artefacts should not be seen as unified objects at all but rather as collections of things (Ayers 1991, Chap. 21). To continue this line of thought would bring in the ontological status of entities like teams, committees, governments, but arguably also of natural entities like ecosystems.

This would connect the philosophical examination of artefacts to the notion of a system, a notion that is looked upon with extreme suspicion since the demise of general systems theory without, as seems to be the general verdict, leaving to posterity anything worth keeping. To argue that this is much too harsh a view is long overdue, not only in view of the continued use of the term in the sciences but even more so in view of its role in technology and engineering. The discipline of systems engineering seems to be here to stay and a gradual shift of emphasis from artefacts to systems and, most recently, sociotechnical systems, is clearly noticeable.[10] Renewed philosophical interest in the term will contribute to much-desired clarification of the many conceptual and practical issues involved.

3.5.2 *The Concept of Design*

Although the concept of design intuitively has ties with several concepts central to philosophy – e.g. rationality, planning, method –, design itself seems not to be recognized as something deserving to be thematicized within philosophy. The 'Dual Nature' programme did much to emphasize the relevance of studying design for the philosophy of technology, and did so primarily in support of its work on (technical) artefacts and (technical) functions (Houkes and Vermaas 2010; Kroes 2012). For design in relation to technology generally, the founding contribution is arguably Herbert Simon's (1969) reference to a 'science of design' and characterization of its content. Simon, however, was not a philosopher but a scientist; his worry was to

---

[10] See, for instance, (de Weck et al. 2011).

separate the design aspect of engineering from the science aspect of engineering, distinguishing between ‘engineering science’ and ‘the science of engineering’, where the latter coincided with design, conceived as an activity undertaken in a systematic, methodical way to a level similar to the level where scientific research operates. Being primarily a scientist, even if a philosophically inclined one, Simon lacked the conceptual acumen that a philosopher could have contributed, and his list of the components of this ‘science of design’ was quite heterogeneous. Arguably his presentation should be read first of all as an invitation to start to reflect philosophically on this discipline which, as he said, was “emerging at the present time” (ibid., p. 58), similar to the way philosophy of science has worked on clarifying the scientific method (up to almost annihilating it, for that matter). Since then, however, not much has been added to Simon’s sketch of a programme. The heterogeneity of its components matches the heterogeneity of engineering conceptions of design; a case in point is Vincenti’s ‘anatomy of engineering design knowledge’ (1990, Chap. 7).

Notwithstanding the introduction of a ‘science of design’ as something that should be contrasted to scientific research, developments with respect to both run parallel. The demise of all attempts to ground rational scientific method on a rigid logical basis – if not standard deductive logic then perhaps something close enough *qua* rigidity – have brought on an understanding of theory choice in science as akin to decision-making, characterized by the inevitability of making trade-offs between various requirements and desiderata, similar to how engineers are used to approach concept choice in design. Not surprisingly, therefore, similar results have been claimed concerning the problems associated with the view that trading-off can be done in a rational way (Franssen 2005; Okasha 2011). Here the development of philosophy of technology will obviously include a charting of the border area with philosophy of science and may be expected to reveal similarities as well as dissimilarities. Indeed one of us has recently argued how vast this border area is where the two are deeply entangled (Franssen 2015).

In one respect a clear difference between the two fields has to be noted: for design, far less work has been put into trying to capture the logical form of the reasoning that goes on in it. It is, for instance, remarkable that the difficulties presented by Von Wright in the 1960s concerning the structure of practical reasoning and the logical connections between the types of statements involved have not been resolved. Niiniluoto (1993) has pointed out the centrality of these questions to the idea of a ‘science of design’. Concerning these matters, the philosophy of technology arguably can profit much from what goes on in logic and, increasingly, technical off-spring areas like artificial intelligence, where many ‘logics of action’ have been developed. Examples of what introducing logical expertise to well-known problem areas in technology has to offer are the analysis of means-ends reasoning (Hughes et al. 2007) and of the assignment of responsibility in situations where ‘many hands’ contribute to an end situation (de Lima and Royakkers 2015).

3.5.3 The Notion of Use

The notion of use, finally, seems so beguilingly innocuous that one may wonder whether it deserves special scrutiny and how it could serve to characterize the philosophy of any practice. Accordingly, it has not been problematized in philosophy – nor, for that matter, in technology and its philosophy. However, failure to recognize that the concept is of central important to its own field and merits considerable more scrutiny than it has received until now may exactly be a key indicator of philosophy of technology's immaturity as a branch of philosophical inquiry. And indeed it is with respect to the notion of 'use' that we end up saying the most and in terms of it that we currently see the most promising steps to be taken on the road to developing philosophy of technology into a serious branch of philosophy.

The process will include expelling spirits from the past. In one of the very few publications dedicated to 'use', Preston (1998) claims that 'using something' has exclusively been approached as indicating a (scientific) type of behaviour by an agent but that all attempts to single out the type have failed; she proposes an alternative approach based on Heidegger's account of the 'readyness-at-hand' of common tools in *Sein und Zeit*. This, however, definitely looks like a false start against the background of the enormous and still-expanding variety in the technological manipulation and control of the natural as well as the social world. Some steps toward an account of use that does justice to this variety and to using non-technical things in non-technical situations have recently been made by one of us in an attempt to extend the research programme 'The dual nature of technical artefacts' to systems rather than 'single-device' artefacts (Franssen 2014). It approaches use not as a type of behaviour but as a relation existing in a system of one or more entities that stand in specific relations to each other. The focus is not on the properties of or behaviour of particular using agents or particular objects used, but on how entities instantiate the specific relations or 'roles' of user, used instrument and used-for object. This brings out another way in which the notion of system is relevant to philosophy of technology, as already noted above in Sect. 3.5.1.

A further issue is how the notion of an artefact's use relates to its 'function'. A key problem for understanding their relation is the notorious re-appropriability of artefacts, which indicates the possible divergence of intended and actual use. A church can become a disco, a social housing unit be used to detain political prisoners, and a hammer can be used as a paper weight.[11] This has proved to be a major obstacle for clarifying the notion of function throughout technology, leading to its being replaced by 'use', as suggested in a recent presentation by Christoph

---

[11] Recent and extensive treatments of the divergence of intended and actual use of artefacts from a philosophy-of-technology perspective, defending opposing positions, can be found in Houkes and Vermaas (2010) and Preston (2013). The first two of our examples are discussed by Sauchelli (2012) and Priemus and Kroes (2008), resp.; the third is ubiquitous in the literature on the functions of artefacts.

Baumberger for the case of architecture.[12] This, however, as Baumberger argues, is a false start, since 'function' in architecture relates to an intended property the architect or designer imbues the building with, whereas 'use' is an acquired property that a building gains at the hands of users, and the frequent non-coincidence of the two (e.g., when a planner fails to anticipate a user's intention) indicates that it is mistaken to assimilate 'function' to 'use'. Unlike Baumberger, however, we hold that the notion of 'function' is too confused to do heavy duty work in the philosophy of technology, and thus assign explanatory work to the notion of 'use' instead. In this, we are tacitly reversing a proposal in Crilly (2010, p. 312):

> This article suggests that by thoroughly exploring the concept of function we can consider all uses of artefacts to exploit artefact functions; they just exploit different kinds of function. In this sense, the terms 'use' and 'function' have a similar scope, but people use artefacts, whilst artefacts perform functions. In returning to our example of the car, we would then simply say that using a car for transportation is to exploit one of the car's functions (a technical function), and using a car to express personal values is to exploit one of its other functions (a social function). This is not just a question of semantics, but a distinction that is useful for two reasons: (i) the underlying connection between seemingly remote uses can be revealed; and (ii) well-developed ideas from one domain can be deployed in a domain in which ideas are less well-developed.

Basically, we agree with Crilly that treating *both* 'function' and 'use' as explanatorily primitive would hardly testify to explanatory parsimony, but we disagree with his clause (ii) on where to assign that primacy, although we share his reason (i) that artefactual, social, and aesthetic 'uses' should receive joint, systematic philosophical attention. Concrete results attained in Crilly's proposal or our competitor would not merely testify to the relative success of either but to the agenda shared by the programmes: to clarify how 'function' relates to 'use'.

Since most of the work on clarifying 'use' still has to begin, we shall for the remainder take for granted a pre-theoretic concept of use that ranges over intended and appropriated uses of an artefact (or, for that matter, of an appropriated natural thing). Further, we eschew any attempt to analyse use as a *criterion* for or (sole) *definiens* of 'artefact'. Instead of focusing our philosophical energies on such analytic endeavours, Sect. 3.6 deploys this (pre-theoretic) concept in order to raise questions *about* artefacts: which properties they have (and do not have), and which moral, metaphysical, and epistemic consequences artefacts having these (and just these) properties has for a philosophical account of artefacts, and consequently (if indirectly) for a philosophy of technology. This suffices, in turn, to link our themes (1) to (3). But before we start doing so, we first draw out a few more general theses about artefactual 'use'.

Clearly standards of rationality apply to use: using involves planning. Use is typically use of objects, but (*i*) not necessarily of objects, and, (*ii*) when of objects, not necessarily of artefacts. According to the system conception referred to above,

---

[12] Baumberger, 'Funktion und Gebrauch', presented at the second Forum Architekturwissenschaft, Darmstadt, Germany, November 2015, abstract available at: http://www.architekturwissenschaft.net/pdfs/Abstracts_Forum_Architekturwissenschaft_2015.pdf

using something necessarily involves the recognition of two interfaces in the ‘thing’ to be used. An artefact is an object with features that are intended as interfaces. Use therefore involves *planning* how to manipulate an entity and what this will result in, plus the actual manipulation. The planning part may seem wafer-thin, but is never entirely absent, we would claim, contra Heidegger. The use of a hammer to hammer in a nail, even if it is granted that how to hold the hammer is typically not a topic for deliberation, still always involves a decision on where to place the nail and how to hold it, which we consider integral parts of using a hammer.[13]

It is through the involvement of planning that a general analysis of using things seems also of central importance to ethics. Many specific forms of use have received their own separate treatment in ethics, e.g. the use of persons, the use of software, the use of information. A generalized conception of use could help to show what these different treatments can contribute to each other and to show how each of them could be extended, as we argue in Sect. 3.6. Indeed, some ethicists have recently developed entire normative and meta-ethical frameworks in which the concept of *planning* is taken as primitive, to then explain our most fundamental normative notions. Prominent among such frameworks is the work of Allan Gibbard. He explains the overall intent and starting point of his book *Thinking how to live* as follows (2003, pp. xi–xii):

> I simply start from us as planners, able to think what to do now and in future contingencies, and conducting thought experiments on what to do in plights that are merely hypothetical. From this starting point, familiar normative phenomena emerge: We see how *ought*s ‘supervene’ on natural *is*s, the way that what a person ought to do supervenes on the natural facts of her situation. […] I begin, then, as a naturalist about humanity, about human thinking and planning [.]

More recently, Gibbard expanded the explanatory use to which the primitive appeal to ‘planning’ is put to even more fundamental normative notions. In his Meaning and normativity, he defends (2012, p. 140):

> …the thesis that ought beliefs are states of planning, in a sense, so that beliefs concerning meaning are plans as to what sentences of one’s own language to accept under various suppositions and given various dispositions that one might have.[14]

If Gibbard’s theses on the normativity of (moral) action and belief are correct, philosophical analysis of the notion of *planning* could have momentous consequences not just for a philosophy of technology, but for clarifying foundational issues in practical philosophy more widely – thus generating a central contributory role for philosophy of technology.

---

[13] Philosophers of technology in the later Heideggerian tradition have connected this point to issues of ‘tacit knowledge’ and ‘knowing how’; space does not permit a discussion of these issues here. Within the ‘Dual Nature’ research programme, the notion of a *use plan* – initially termed *user plan* – made the terms ‘use’ and ‘plan’ central, if not almost constitutive of artefacts, but the terms were applied as if unproblematic and the application was not accompanied by an analysis of either using or planning (see Houkes et al. 2002; Houkes and Vermaas 2010).

[14] Defended more fully in Chap. 8 of the same book.

3.6 Metaphysics and Ethics: A Use-Theoretic Individuation of Artefacts

In Sect. 3.5.3, we suggested to relate our core themes of artefact (1) and use (3) by examining (*i*) how the properties of artefacts relate to their intended and non-intended uses, and (*ii*) how this 'relation' impacts (and is impacted by) which properties artefacts have. This, we alleged, would help (*iii*) bring out crucial moral, metaphysical, and epistemic consequences for a philosophical account of artefacts – the type of account that can serve as an adequate position in answer to core theme (1). The present section fills in the detail of this outline. Section 3.6.1 turns to (*i*) and (*ii*) and Sect. 3.6.2 uses the results thereof (and a great deal besides) to deliver a more systematic analysis of 'the use of an artefact'. The closing Sects. 3.6.3, 3.6.4, and 3.6.5 then turn to the 'consequences' (*iii*) this has for ethics, metaphysics, and epistemology. We start by looking more closely at how the properties of artefacts relate to their intended and non-intended uses, and how this impacts (and is impacted by) which properties artefacts have.

3.6.1 Artefacts and Properties

If the range of a thing's possible re-appropriations seems to be wide open, the properties an artefact has (pre-use) set a limit here. One cannot use a sheet of A4 paper as a paperweight, no matter how much one modifies (chemically or otherwise) its existing properties. At the same time, not all pre-use properties a thing has will ever be 'use-relevant'.

Let us call the properties of a thing that a user can recognize as 'interfaces for manipulation' (Franssen 2014) this thing's 'use-relevant properties'. (Typically we think of such interfaces in surface or macroscopic terms, although the onset of nanotechnology has shifted that surface to a considerable depth.) If we regard a given use of an object as an action by an agent ('manipulator') to achieve a set goal, and that goal to be only attainable by the manipulation of an object, we can then regard the object as a 'vehicle' for that use. In that case, we can stipulate that those (and only those) properties of the thing in question are 'use-relevant' that are necessary for the agent to achieve the set goal. These are the thing's use-relevant properties relative to a given use, in its turn relative to a set goal. We can remove this relativity by replacing the given use and set goals with an unbound variable, such that a (given) object's 'use-relevant' properties are those, and only those, that can help an agent manipulating that object attain some or other set goal.

These preliminary constraints on which properties qualify as 'use-relevant' set the opening to the joint systematization of 'use' and 'use-relevant': we ask, in particular, given a thing and its 'use-relevant' properties (where the 'thing' may be a material object but also, say, a sentence), to which uses that thing can be put. But even then we shall not lapse into the philosophical error, or premature ambition, that systematization of this kind affords an analytic definition of 'use', or 'use-relevant'. The only result by way of definition is a tautological one: a thing considered as (or *qua*) artefact will only have properties of two kinds: (*a*) 'use-relevant' properties, and (*b*) the uses to which it can be put. If we conceive of (*b*) as properties of the kind *being used by an agent* S *to do* X *to achieve* Y *in world* w *at* t, then (*b*)-properties are *necessarily extrinsic*. By contrast, (*a*)-properties will be *possibly intrinsic* to their bearer[15] – this depending on whether a given property such as mass is relevant to one of its bearer's possible uses, and whether we champion a physics that considers that property (mass) to be an extrinsic rather than intrinsic property. And this contrast, of necessarily extrinsic to possibly intrinsic, is one of many philosophical results that a systematization of, not just what falls inside the (*a*) and (*b*) range but of the *relation* between these two property types, can unearth. These results then drive further philosophical reflection on our concepts of use and artefact. For instance, it becomes clear that appeal to 'use-relevant' properties by themselves does not serve to delimit the domain of artefacts. A stone pebble on the moon could be appropriated by some use at some point in time (and with minimal alteration), but is hardly an artefact. Natural things differ from artefacts in that the properties they (can) have is not exhausted by properties of type (*a*) and (*b*).

3.6.2 Varieties of 'Use': From Philosophy of Language to Philosophy of Technology

To improve our grasp of the notion of an artefact's 'use-relevant properties', we argue that philosophy of technology can be fruitfully linked to work done in the philosophy of yet another practice, namely language. Linguistic pragmatics focuses on use as a crucial aspect of language, and recent advances in the philosophy of language on the *use* of particular expressions at the sentential and subsentential level are characteristically presented under the guise of various 'use theories of meaning'.[16]

To illustrate, suppose a policeman walks up to you just after you have parked in the wrong spot, and says 'You do not want to park here'. That statement (*S*) has a literal truth-condition, and the policeman (and you) perfectly understand what it is. What is more, both of you understand perfectly well that the policemen is not committed to the truth of that statement in uttering it (it is most likely *literally false* – why would you have parked your car here if you did not want to?), since the statement's point (which the policeman very much *is* committed to) is a different

---

[15] An intrinsic property is a property a thing can have *by itself*, whereas an extrinsic property assumes its bearer to be related to *something other* than itself. While philosophers disagree on how to remove the circularity in this 'definition', they agree that the definition provides a satisfactory informal elucidation of 'intrinsic'.

[16] For a recent overview, see (Bach 2004).

one – you should not park here, and he will write you a ticket if you do not move your car promptly.[17]

To account for such divergences between non- and literal uses of a sentence, Austin (1962) introduced a framework designed to overcome philosophical obscurity attaching to the phrase 'use of a sentence'. Austin deemed that phrase obscure since "'use' is a hopelessly ambiguous or wide word, just as is the word 'meaning'" (p. 100). We follow suit to overcome a similarly 'hopelessly ambiguous' phrase: 'use of an artefact (or of a technology)', and begin by adopting Austin's three varieties in which one and the same sentence *S* can 'mean' something, depending on whether *S* is used, and if so, in what manner (*idem*, pp. 99–109)[18]:

(*a*) The 'locutionary': this is *S*'s meaning as accounted for by (entries in) a standard lexicon (or dictionary) and grammar of the English language. In the example given, the sentence *says* that the addressee ('you') does not want to park in a certain spot, but *S* itself is neutral as to what point (if any) speakers might make in *using* that sentence. Hence, at stage (*a*) *S* has no concrete relation to specific occasions of 'use' yet but is rather seen as purely a 'vehicle' of use.
(*b*) The 'illocutionary': this is the generic *use* to which *S* is put in the broad taxonomy of basic 'types' of 'speech act', namely: an assertion, command, or question. Under this broad taxonomy, the sentence in our example qualifies as an assertion.
(*c*) The 'perlocutionary': the bringing about of things that *a subject manipulating S* can do by using *S*, where the content of these 'things' typically extends beyond what we find in categories (*a*) and (*b*). In the present example, the policeman discourages *S* from parking by issuing a subtle warning: the fact that the policeman does so is well conveyed by *his* use of *S* in the present context, but does not supervene on the properties of *S* we have isolated so far (*a*, *b*).

In the sequel, category (*c*) shall occupy us most. This should surprise readers little, in that Sect. 3.4.1 already stated that, for us, technology (or art for that matter) is fundamentally not about making things but about making things happen – in a way that crucially involves objects or 'vehicles'. This is what Austin's category (*c*) is meant to elucidate. Philosophers have recognized the wide variety of 'use' *S* receives within category (*c*) alone, and offered preliminary taxonomies. Bach (2004, p. 467) mentions four categories, such as (*i*) statement (the utterer *U* intends the hearer *H* to pick up a belief reported in statement S), (*ii*) request (*U* desires *H* to do as requested in *S*), (*iii*) promise (*U* intends *H* to belief that *U* has a sincere intention to do *D*), and (*iv*) apology (*U*'s uttering *S* enacts and expresses *U*'s regret for and intent to

---

[17] See (Koller 2015, p. 31).

[18] Our adoption of Austin's framework departs in one point, concerning category (*a*): Austin thought (where we presently do not) that sentence meaning was already an abstraction of sentence use, namely of "making a statement" (*idem*, p. 1 fn. 1). This has the confusing result that Austin thinks of (*a*) as already a type of speech act, with the consequence that no such speech act can occur in the absence of type (*b*) speech acts: "To perform a locutionary act is […] *eo ipso* to perform an *illocutionary* act" (*idem*, p. 98).

compensate *H* for a past action of *U*). Bach (ibid.) subdivides each of (*i*)-(*iv*) into further varieties: in (*ii*), or 'directives', we find "admonishing, advising, asking, begging, dismissing, excusing, forbidding, instructing, ordering, permitting, requesting, requiring, suggesting, urging, warning", and in (*iii*), 'acknowledgments', "apologizing, condoling, congratulating, greeting, thanking, accepting (acknowledging an acknowledgment)".

In accumulation, such taxonomies along Austin's lines illustrate, firstly, the diversity of 'use' and use manifestations in general (not limited to specific sentences), and secondly, interim success at systematizing the very notion of 'use' itself without betraying exactly that diversity of use-manifestations. The taxonomy highlights the *flexibility and systematicity* of the *relation* between the properties that S has as a type-sentence of the English language prior to its use and the particular uses to which it is put.[19] In the above example, not *any* old (type) sentence can be used by the policeman to issue a warning: certain background conventions, shared and understood by utterer and addressee, have to be in place, and these conventions are *sensitive* to the properties the linguistic vehicle *S* has by itself.

These points cry out for re-appropriation and extension in the philosophy of *non-linguistic artefacts*.[20] We may think of phonetic strings at sentential length as human artefacts with sentential structure. Speakers produce and manipulate such artefacts to specific ends, by circulating and sharing them. It seems in principle possible that a wide range of *non*-linguistic human artefacts, especially artworks, can help perform a similar range of functions. Figurative painting, for instance, or drama (itself a string of highly contrived sentential strings uttered in a specifically arranged context), can easily serve complex perlocutionary ends. And so can architectural buildings if we go along with Wittgenstein's suggestion to understand architecture 'as gesture'.[21] Wittgenstein's idea, it seems, was to begin with a *generic theory of gesture* to be applied to speech acts and buildings. This approach enabled him to both show the parallels of the two types of artefacts, and at the same time allow for the idiosyncrasies of each. In gestures of either kind, the agent's intention *is expressed*, and *meant to be recognized*, in her act, where in "(mere) purposive movement" (such as walking down a street) one's intent and volition receive no overt expression.[22] This observation re-connects us to – indeed, may lend structure and viability of an actual research program to – Mitcham's proposal to consider (some) artefactual objects as 'expressions of human volition'.

---

[19] This is sometimes called 'sentence meaning' (to contrast 'speaker's meaning'), or (a sentence's) 'literal meaning'. This taxonomy has been hotly disputed in recent years, but thankfully little of that debate impacts the extremely basic points we draw on here.

[20] Thomasson ends her (2014) contribution to a volume on artefact kinds precisely with the claim, or rather suggestion, that there are commonalities between technology and language. Language here features as another practice, to which technology could be compared and with which it could be contrasted (*mutatis mutandis* for the philosophies of technology and language). This is certainly how we intend readers to view the present section.

[21] Wittgenstein (1998, p. 50).

[22] Wittgenstein (ibid.). In merely walking or drinking a cup of tea, the agent would not (have to) seek to have her intentions recognized in the act itself.

The approach to understanding artefactual use sketched here, drawing on Austin and Wittgenstein, is not novel. Its starting point is observing the similarity in use of non- to linguistic artefacts, as observed in Wittgenstein (1953).[23] Krippendorff (2005) has thought this 'similarity' sufficient to motivate a 'semantic turn' in the philosophy of technology; one of his reviewers rather thought it motivated a 'pragmatic turn'. While we share with Krippendorff the *starting motivation* of such a project – to enlist Wittgenstein's use theory of meaning to analyse artefact use –, we differ sharply from him in the theoretic tools to enlist in the program's execution. By contemporary standards, Wittgenstein's work on pragmatics (driving Krippendorff's project) is (1) too "anti-theoretical" in its adversity to systematization, and (2) shows theoretic commitments – such as the inseparability of semantics from pragmatics – that are no longer deemed tenable (Bach 2004, p. 463). In both regards, we side with the tradition in pragmatics inaugurated by Austin and Grice, as per our framework (*a*) to (*c*) above, and reject (1) and (2).

A use-theoretic approach to understanding technical artefacts – by extension, to all central notions in the philosophy of technology that can be defined in artefactual terms –, if it is to be systematic and generic, has to find its place in a spectrum of rational behaviour ranging from the *extremely codified* and convention-based (such as speech acts) to more free-form areas (such as, kicks under a table). Artefacts successfully located on that spectrum (depending on artefact type, we may have to place them on different points in that spectrum) not only enable us to bring in nuanced theories of artefactual usage, but relate those theories in more concrete terms to extant philosophical debates on: language use, rational behaviour, practical rationality, planning. By multiplying these points of subdisciplinary connection, we raise standards of adequacy expected of the analysis of distinct artefacts' use.

This, in turn, is crucial in that the approach here sketched also constrains the moral analysis of artefactual usage. To see this, we need to posit artefactual analogues to Austin's categories (*a*) to (*c*) above. For each artefact (type) *x*, we have to determine the artefact's *relation to* use, even if (as in Austin's scheme) the first category demarcates the artefact's properties *prior* to actual usage:

(*a**) the quasi-locutionary: the use-relevant properties of *x*
(*b**) the quasi-illocutionary: the 'narrow' use of *x* that stays constant across different practical contexts, and
(*c**) the quasi-perlocutionary: what one's specific narrow use (*b**) of *x* is managing to accomplish. (*NB* As we shall see, recursions on *c** are possible too. Hence, the present scheme can be continued, as can the previous one, (*a*)–(*c*), although that is not normally considered.)

For instance, a gun's *a**-properties include its weight, stability, etc., enabling it to function in manner *b**, and its *b**-property is its ability to compel a projectile at a certain distance and velocity. Put differently, *a** constitute the gun's capacity to

---

[23] "Wittgenstein insisted on the importance of understanding meaning as use and not separated from practice. Language does not represent artefacts, but is itself an artefact we use when we participate in intertwined language-games." (Ehn 2007, p. 56)

compel a projectile; *b** is the disposition of firing the projectile. In established jargon, *a** is the categorical base for the gun's dispositional (use) property *b**.[24] The gun's *c**-property denotes the various 'uses' to which the gun can be put *insofar as* (*and only insofar as*) it has the properties denoted by *b**. Thus, if used to potentially harm, maim, or kill humans, the gun will be classified at the *c**-level as a weapon. But this is not the end, for the schema 'putting *x* to use in light of *x*'s properties *P*' allows instantiating in the *P*-variable not only an artefact *x*'s – say, the gun's – *a**- and *b**-properties, but endless recursion on *c**, and recursions on recursions of *c**, indicated by *d**, *e**, and so forth. That is, a gun used as a weapon may be used (*d**-1) by a robber to intimidate and if necessary disable a shop keeper; the same gun may be used (*d**-2) on a policeman's belt to deter a burglar; it may be used, just as well, (*d**-3) as an exhibit in a museum piece. Finally, the gun's use in its incarnations *d**-1 to *d**-3 may figure in larger uses, such as a 90-min crime show displaying scenes in which the same gun is used to all three of these effects, and so on. As we move from narrow (*b**) to more complex (*d**) recursions on the same gun's use, the question arises to which degree such further uses actually draw on features beyond the gun's 'use-relevant' ones it has by itself. One could obviously not 'shoot' (no pun intended) a TV show in which at some moment a gun is going to be fired without having a great deal besides the gun at one's disposal and among one's assets: at this point, the gun and its use have simply become components of a much larger picture. (In language use, '*d**' would mark the stage of a sentence featuring in a movie script or stage performance.)

Still, the fact remains that across such increasing degrees of recursion, the gun's properties by itself stay morally relevant – that is, stay 'use-relevant' to specified uses and ends. We can even introduce that observance of that 'fact' is now mandatory to any moral appraisal of (especially 'new' or 'innovative' technologies): such appraisal needs to demonstrate (*i*) its being grounded in a (demonstrated) understanding of the technology's 'use-relevant' properties, and (*ii*) needs to specify the precise recipient of moral appraisal (see further Sect. 3.6.3). Here are two hypothetical examples of violating (*i*) and (*ii*):

> Imagine a paper written on the ethics of hybrid cars. At face value, the paper presents its claims as being about the moral drawbacks and advantages of cars designed with a hybrid engine. However, on closer inspection (neg-*i*) the paper fails to mention the morally salient properties of that engine, and (neg-*ii*) it turns out the paper's moral key claims all pertain to issues of distributive fairness in the setting up and administration of smart grids supplying the electric energy of hybrid cars. If the analysis had failed on points (*i*) and (*ii*) and the paper had concluded that the *design* of such cars needs to be 'sensitive' to moral values, the conclusion would fail to be compelling.[25]
>
> Imagine a paper written on the ethics of drone warfare. As before, (neg-*i*) readers learn no specifics that specify the use-relevant properties of drones and the uses to which they are put, other than a vague notion that drones are 'robots that kill'. Then, (neg-*ii*) the paper

---

[24] See Shoemaker (1984, pp. 206–260), and Campbell (2002, pp. 235–253).

[25] The present example can be replicated to many other types of artefacts – such as healthcare robots –, as soon as the artefactual specifics of such artefacts and their conditions of use fail to receive sufficient analysis.

> raises issues about the immorality of killing humans in a certain political conflict C that at one point may or may not have featured drones, or where C's political masterminds entertained the idea of employing drones. (Neg-*i*) shows that the paper's main moral considerations fail to engage the artefact in question – by neglecting analysis of its 'use-relevant' properties – and instead obfuscates the issue by moralizing a political conflict C, where C's main morally relevant features (as analysed in the paper) would stay constant regardless of whether or not drones are employed.

Apart from acting as a *methodological constraint* on socio-moral technology assessment, our framework's insistence to relate uses to (a proper individuation and analysis of) 'use-relevant' properties has three larger consequences for a philosophically systematic understanding of technical artefacts – consequences in ethics, metaphysics, and epistemology.

3.6.3 Ethics

As we just saw, an individuation of technical artefacts in terms of usage (a^*)–(c^*) has consequences for a unified theory of technical artefacts in the domain of ethics. A further connection, of the metaphysics (and uses) of artefacts, to a live issue in contemporary ethics, termed 'recursive ethics,'[26] transpires here. The position at the issue's centre involves (at least) two core claims. One, recursion seems to underlie a claimed analogy between language and ethics. For instance, the meaning of increasingly complex sentences formed by sentential connectives (like 'and' and 'or') can be specified recursively, namely in terms of the complex sentence's atomic sentential components *and* the meaning of the connectives. Just so, it is thought, can we specify the moral value of increasingly complex acts that interconnect analogously 'atomic' components. For instance, an offer to kill someone else or a promise to do so in part inherit the immorality of the act of killing, even though they do not (yet) constitute such an act. Similarly, an artefact's ability to entice – like a promise – a certain immoral act can thereby inherit a degree of that immorality. Naturally, the term 'entice' has to receive a precise gloss, one that is maximally specific on the factors we outlined above – the artefact's properties it has by itself, and the *relation* of those properties to specific (here, enticed) use.[27]

Analysis of the term 'entice', and construction of the ethical theory this analysis will inform so as to yield systematic appraisal of the moral status and value of technical artefacts, has to, finally, fully draw on a third factor we have now developed: the recursive nature of artefactual use itself, ranging from level b^* to d^* and beyond. If, say, we can pin moral value on x's use b^*, then a systematic theory of artefacts' moral appraisal has to examine and explain how the moral value of x's use d^* relates

---

[26] See Alfano and Loeb (2014) and Hurka (2011).

[27] It seems to us that none of the standard theories currently on the market – action schemes, mediation, etc. – manage to shed much light on the matter. In the final instance, the 'entice' remains unanalysed, and thus fails to explain what we wanted to know: the nature and degree of (im)moral 'inheritance'.

to the moral value of *x*'s use *b** if and only if *d** is a recursion on *b**. This is *not* to say that the vague relation indicated here by 'relates to' can be instantiated by a precise function or computation with *b** and *d** as sole input. For, firstly, *d** may comprise a lot of further elements that determine its moral value, aside from *b**, and secondly, the relation between the two moral values could be rather different than the relation between the two uses.[28] All we claim here is this: an explanation of *d**'s moral value has to *account for* the moral value of *b** if and only if *d** is a recursion on *b**. But *that* is just reporting *a* relation between the two (one that needs attention in a moral analysis of *d**) rather than the affirmation of a *distinct* (already analysed) relation between the two.

Even then, one of the most obvious advantages of this framework is to retain systematicity under diversity of use. We move away, for instance, not only from the philosophically unwarranted simplification of construing technical artefacts as morally neutral, but also from their equally simplified construal as *always fundamentally cooperative* and *benign toward* '*the user*', and can explore non-benign relations of use, as well as transformations of morally benign into non-benign forms of use.

3.6.4 Metaphysics

The approach sketched above mutually constrains the metaphysical individuation of artefacts as objects and the analysis of 'use'. And this has consequences when approaching debates in philosophy about the individuations of artefacts (seemingly) unconstrained by facts of use. The following is an example.

Amie Thomasson says (2007) that attention to technical artefacts may revise how we do metaphysics of artefacts more generally. Elsewhere (2004), she argued for a similar impact of attention to artistic artefacts. Perhaps that is so, but Thomasson's hope remains relatively vague. But once we have a working metaphysics of technical artefacts on the table, we can be clearer concerning how much and what precisely of general metaphysics is relevant. This ties in with a recent paper by Karen Bennett (2009). Bennett's overriding concern is to demarcate substantive from 'purely verbal' disputes in metaphysics; her paper's relevance to current concerns is wholly borne out by her use of a technical artefact – a toaster – as her primary example.

> *Composition*: In my kitchen there are some physical particles which together *compose* my toaster (are arranged such that they form a toaster, an artefact which can toast bread). Some people believe that there are only those particles ('simples') but nothing more, whereas others believe that there is also *the toaster*, an entity formed by those simples.

---

[28] For example, even if *b** is morally neutral, that of *d** could be extremely negative (call this *moral value amplification*); and the reverse is conceivable too, where *d** is morally neutral even though *b** is morally negative (call this *moral value neutralization*). So, putting a dangerous weapon behind a secure glass display neutralizes its potentially harming uses, analogous to how some adjectival modifiers like 'allegedly' neutralize the ascriptive content and moral value of what they qualify, such as 'is a murderer'.

In Bennett's terms, this dispute is not substantive because even those who believe that there are only particles (and no 'toaster') would never say "Honey, revise your breakfast plans: there's no toaster in the kitchen" (p. 58). That is, it seems unlikely that two people (whether inside or outside the present metaphysical dispute) would ever disagree 'what it takes for something to be' a toaster or a table. Even a 'nihilist' about toasters does not reject countenancing toasters because whatever is in the kitchen just 'does not cut it' (as to what it takes *to be* a toaster). Rather, he thinks *nothing* in the world qualifies for being a toaster, precisely because he is hesitant to *countenance* toasters. All parties to this dispute actually agree *what the world is like* (here: 'there are simples arranged to form a toaster in my kitchen, so that bread can be toasted'), but disagree what these agreed on claims *give us license to infer*: can we infer that *there is* a thing called 'the toaster' in the kitchen? It seems a moot point – whether or not there is such a thing, we can have toast for breakfast. If, as we have argued above, the defining properties of technical artefacts are exhausted by their 'use-relevant' properties, and the latter stay constant across divergent ontologies, it follows that (at least some) foundational disputes on the ontological nature of 'toasters' carry limited relevance for a metaphysics of technical artefacts.

3.6.5 Epistemology

An issue not yet discussed in the above is the epistemology of (making or using) artefacts. Debates on delineating 'technical (or technological) knowledge', esp. as pertaining to the *making* of artefacts, frequently insist on the type-distinctness of such knowledge to scientific knowledge, but typically argue for such distinctness on the grounds of a distinction between knowledge-that and knowledge-how that has recently come under much philosophical fire. Without wishing to mine such recent debates within pure epistemology (i.e., foundational analytic philosophy), let alone take a stance on what is a debate in mid-progress, we think that the present issue can *avoid* reliance on such controversial notions – and achieve the same, if not more, clarity on our lead question by going for a 'use-theoretic' framework.

To begin with, we assume that the notion of 'making' relevant to the production of technical artefacts has such artefacts as the causally distinct upshot of the process of making. In the 'making' of certain artistic artefacts, especially those related to performance of song and dance, this is arguably not the case.

We thus restrict 'making' to a relation defined in terms of *causality* and *token object*. Making is the production of specific objects, and once the causal relation is understood, the process can be defined in terms of the object the process gives rise to. At this point, we have to ward off a physicalistically reductive reading of objects produced. Rather, the making of technical objects is at the very least also a teleological process, where a certain end goal *directs and shapes* the process leading up to that object's gestation. In order to not mystify this point, we load the 'goal-directness' not into the causal process, but into its causal outcome – the artefact produced in the 'making'. We can do this by individuating our artefact with respect to such goals. As such goals are defined by the object's intended use, we have arrived at the object's 'use-relevant' properties. That is, insofar specific uses require vehicles of usage (things that help us perform discrete uses), artefactual making seems to be *of the vehicle*, which is individuated with respect to its use-relevant properties. An engineer would *make* vehicles of uses, not uses themselves. (Compare: in extending your hand you might offer *a handshake*, but it would hardly be felicitous to describe this as the recipient 'using' your hand.)

The only thing that remains to be done is to link our building blocks – making as a (special) type of process, and that process as one guided by a teleologically defined object – to the notion of knowledge. We can thus define knowledge of engineering as follows:

(**KE**) Knowledge of engineering $=_{\text{df.}}$ knowledge of what it is that enables the engineer *to endow* the artefact with use-relevant properties.

This proposal is of course rather obscure, since one of its key terms, '(to) endow' is mysterious. So we revert to a slight re-formulation:

(**KE***) Knowledge of engineering $=_{\text{df.}}$ knowledge of what it is that enables the engineer *to see to it that* the artefact has use-relevant properties *P*.

To be sure, the italicized phrase in KE* still designates an intensional operator, but one we can understand with relative ease. For, with KE* in place, we can ask, and resolve with relative precision, a philosophical question frequently faced by philosophers of technology: is knowledge of engineering a practical ability, to be analysed as 'knowledge how'? While the literature frequently seems to lean towards an affirmative on this matter,[29] our reconstruction of artefacts' epistemic properties suggests otherwise. For KE* is structurally analogous to:

(**KW**) Knowledge of keeping your children warm in winter $=_{\text{df.}}$ knowledge of how your children have to be attired to keep warm when the day temperature is in a certain range (say, below 5 or 10 °C).

But the knowledge mentioned in KW strikes us as nothing mysterious, or inevitably non-propositional. Rather, one needs to know about proper attire, relative to the insulating properties of clothes, and how closely they match the physical requirements and overall shape of one's children's bodies. And this, in a nutshell, fits classic means-ends reasoning that Aristotle (*de Motu Animalium* 7) thought to typify the notion of 'craft', of technical making, and was at work in such tasks as the making of a coat.

Naturally, KE* might not contain all the answers we want. It says nothing about how to select the properties denoted by the plural variable *P*. It seems that an artefact's intended use may help somewhat.

Further, KE* might say *more* than can be ultimately vindicated, especially its attendant claim that analysis of 'sees to it that' reductively eliminate a thicker notion of 'design' introduced in Sect. 3.4. One would perhaps rather expect that elaborating

---

[29] See for instance (Houkes 2009) and some suggestion in (Kroes 2014, p. 11).

our key theme (2), the concept of design, will reveal how to best see the *relations* between various other notions introduced since then – making, practical knowledge, practical action, planning – and to make clear to what extent a reductive approach to any or all of them can be expected to be successful.

3.7 Summary and Conclusion

In summary, by pointing the way to a 'use theory' of artefactual objects, we show some of the benefits that might be reaped from connecting the metaphysics of technical artefacts to a 'use theory of meaning' and from connecting the ethics of those artefacts to some core features at work in an 'ethics of recursion'.

And this, in turn, illustrates the essay's larger claims, claims it sketched and in part illustrated. Progress in, and beyond, the 'empirical turn' in especially analytic philosophy of technology hinges (at least) on two components. One, contributions to the field have to be developed with much greater sensitivity to *systematicity*. While the field may not realistically reach a state of a 'unified science', the field can only mature if it leaves behind the mode of the 'perennial freshman', where every new contribution is developed without testing its consistency with established results reached on logically and conceptually related questions. Two, 'progress' will likely require much greater and more focused reliance on advances in foundational philosophy, above all metaphysics and the philosophy of language. Precisely to accentuate what renders technology, and technical artefacts, distinct from (the products of) other human practices, philosophers of technology can profit immensely by focused consultation of philosophical results attained in the philosophies of language and the arts, among others.

Finally, the two components we have shown to be conducive of 'progress' arguably depend on one another. Only by enlisting 'foundational' philosophy can we bring a degree of systematicity to contemporary analytic philosophy of technology, and help secure it the bright future the empirical turn intended for it.

Acknowledgements We thank Peter Kroes, Mark Coeckelbergh, and Pieter Vermaas for comments on earlier drafts; and Mark Alfano for a critical suggestion on what became Sect. 3.6.3.

References

Alfano, M., & Loeb, D. (2014). Experimental moral philosophy. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Summer 2014 edition), http://plato.stanford.edu/archives/sum2014/entries/experimental-moral/.

Austin, J. L. (1962). *How to do things with words.* Oxford: Oxford University Press.

Ayers, M. (1991). *Locke. Volume II. Ontology*. New York: Routledge.

Bach, K. (2004). Pragmatics and the philosophy of language. In L.R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 463–487). Oxford: Blackwell.

Bennett, K. (2009). Composition, colocation, and metaontology. In D. Chalmers (Ed.), *Metametaphysics: New essays on the foundations of ontology* (pp. 38–76). Oxford: Oxford University Press.

Berg Olsen, J. K., Pedersen, S. A., & Hendricks, V. F. (Eds.). (2009). *A companion to the philosophy of technology*. New York: Wiley-Blackwell.

Campbell, J. (2002). *Reference and consciousness*. Oxford: Oxford University Press.

Coeckelbergh, M. (2014). Moral craftsmanship. In S. Moran, D. Cropley, & J. Kaufman (Eds.), *The ethics of creativity* (pp. 46–61). Basingstoke: Palgrave Macmillan.

Crilly, N. (2010). The roles that artefacts play: Technical, social and aesthetic functions. *Design Studies, 31*, 311–344.

Davies, D. (2004). *Art as performance*. Oxford: Blackwell.

de Weck, O. L., Roos, D., & Magee, C. L. (2011). *Engineering systems: Meeting human needs in a complex technological world*. Cambridge, MA: MIT Press.

Dipert, R. (1993). *Artefacts, art works and agency*. Philadelphia: Temple University Press.

Ehn, P. (2007). Review of Krippendorff 2005. *Artifact, 1*, 56–59.

Franssen, M. (2005). Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. *Research in Engineering Design, 16*, 42–56.

Franssen, M. (2014). Modelling systems in technology as instrumental systems. In L. Magnani (Ed.), *Model-based reasoning in science and technology: Theoretical and cognitive issues* (pp. 543–562). Heidelberg: Springer.

Franssen, M. (2015). Philosophy of science and philosophy of technology: One or two philosophies of one or two objects? In S. O. Hansson (Ed.), *The role of technology in science: Philosophical perspectives* (pp. 235–258). Dordrecht: Springer.

Franssen, M., Lokhorst, G. J., & van de Poel, I. (2013). Technology, philosophy of. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Winter 2013 edition), http://plato.stanford.edu/archives/win2013/entries/technology/.

Franssen, M., Kroes, P., Reydon, T. A. C., & Vermaas, P. E. (Eds.). (2014). *Artefact kinds: Ontology and the human-made world*. Cham: Springer.

Gibbard, A. (2003). *Thinking how to live*. Cambridge, MA: Harvard University Press.

Gibbard, A. (2012). *Meaning and normativity*. Oxford: Oxford University Press.

Hilpinen, R. (1992). Artifacts and works of art. *Theoria, 58*, 58–82.

Hilpinen, R. (1993). Authors and artifacts. *Proceedings of the Aristotelian Society, 93*, 155–178.

Houkes, W. (2009). The nature of technological knowledge. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 309–350). Amsterdam: North-Holland.

Houkes, W., & Meijers, A. (2006). The ontology of artefacts: The hard problem. *Studies in History and Philosophy of Science, 37*, 118–131.

Houkes, W., & Vermaas, P. E. (2010). *Technical functions: On the use and design of artefacts*. Dordrecht: Springer.

Houkes, W., Vermaas, P. E., Dorst, C. H., & de Vries, M. J. (2002). Design and use as plans: An action-theoretical account. *Design Studies, 23*, 303–320.

Hughes, J. L., Kroes, P., & Zwart, S. D. (2007). A semantics for means-end relations. *Synthese, 158*, 207–231.

Hurka, T. (2011). *Drawing morals: Essays in ethical theory*. New York: Oxford University Press.

Illies, C., & Ray, N. (2016). An aesthetic deontology: Accessible beauty as a fundamental obligation of architecture. *Architecture Philosophy, 2*, 83–103.

Koller, S. (2015). *The birth of ethics from the spirit of tectonics*. Doctoral dissertation, Delft University of Technology.

Kranzberg, M. (Ed.) (1966). *Technology and culture* vol. 7 no. 3, Toward a philosophy of technology, 301–390.

Krippendorff, K. (2005). *The semantic turn: A new foundation for design*. London: Routledge.

Kroes, P. (2012). *Technical artefacts: Creations of mind and matter*. Dordrecht: Springer.

Kroes, P. (2014). Knowledge and the creation of physical phenomena and technical artefacts. In P. Schroeder-Heister, W. Hodges, G. Heinzmann, & P. E. Bour (Eds.), *Logic, methodology and philosophy of science, proceedings of the 14th international congress (Nancy): Logic and science facing the new technologies* (pp. 385–397). London: College Publications.

Kroes, P., & Meijers, A. (2000a). Guest editor's preface. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (p. xv). Amsterdam: JAI/Elsevier.

Kroes, P., & Meijers, A. (2000b). Introduction: A discipline in search of its identity. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (pp. xvii–xxxv). Amsterdam: JAI/Elsevier.

Krohs, U. (2009). Functions as based on a concept of general design. *Synthese, 166*, 69–89.

Lima, T. de, & Royakkers, L. (2015). A formalization of moral responsibility and the problem of many hands. In I. van de Poel, L. Royakkers, S. D. Zwart, *Moral responsibility and the problem of many hands* (pp. 93–129). New York: Routledge.

Meijers, A. (2009). General introduction. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 1–19). Amsterdam: North-Holland.

Mitcham, C. (1994). *Thinking through technology: The path between engineering and philosophy*. Chicago: University of Chicago Press.

Mitcham, C. (2002). Do artefacts have dual natures? Two points of commentary on the Delft project. *Technè, 6*, 9–12.

Niiniluoto, I. (1993). The aim and structure of applied research. *Erkenntnis, 38*, 1–21.

Okasha, S. (2011). Theory choice and social choice: Kuhn versus Arrow. *Mind, 120*, 83–115.

Preston, B. (1998). Cognition and tool use. *Mind and Language, 13*, 513–547.

Preston, B. (2013). *A philosophy of material culture: Action, function, and mind*. New York: Routledge.

Priemus, H., & Kroes, P. (2008). Technical artefacts as physical and social constructions: The case of Cité de la Muette. *Housing Studies, 23*, 717–736.

Sauchelli, A. (2012). Functional beauty, architecture and morality: A beautiful Konzentrationslager? *Philosophical Quarterly, 62*, 128–147.

Shields, C. (1999). *Order in multiplicity*. Oxford: Oxford University Press.

Shoemaker, S. (1984). *Identity, cause, and mind*. New York: Cambridge University Press.

Simon, H. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Thomasson, A. L. (2004). The ontology of art. In P. Kivy (Ed.), *The Blackwell guide to aesthetics* (pp. 78–92). Oxford: Blackwell.

Thomasson, A. L. (2007). Artifacts and human concepts. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 52–73). Oxford: Oxford University Press.

Thomasson, A.L. (2014). Public artifacts, intentions, and norms. In M. Franssen, P. Kroes, T.A.C. Reydon, P.E. Vermaas (Eds.), *Artefact kinds: Ontology and the human-made world* (pp. 45–62). Cham: Springer.

Vincenti, W. (1990). *What engineers know and how they know it*. Baltimore: Johns Hopkins University Press.

Wiggins, D. (1980). *Sameness and substance*. Oxford: Blackwell.

Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.

Wittgenstein, L. (1998). *Culture and value*. (Ed.) G.H. von Wright. Revised edition. Oxford: Blackwell.

Maarten Franssen M.S. in Theoretical Physics, M.A. in History and Ph.D. in Philosophy, all from the University of Amsterdam. Currently Associate Professor with the Philosophy Department of Delft University of Technology, the Netherlands, and engaging with foundational and methodological issues in the philosophy of science and technology. His specific research interests are the nature of normativity in relation to artefacts and their use, the metaphysics of artefacts, instrumental and sociotechnical systems as constitutive of technology, and the conception of design as decision-making and its problems.

Stefan Koller BA.Hons. BPhil. Oxon., PhD ***cum laude*** TU Delft with ***The Birth of Ethics from the Spirit of Tectonics***. Gibbs Prize in Philosophy, Duns Scotus Prize in Medieval Philosophy, Exhibitioner Award, Scatcherd European Award, and more, all from the University of Oxford. Business Group Member of the ***International Society for the Philosophy of Architecture***. Co-founder and co-editor of ***Architecture Philosophy***, an international journal to simultaneously pursue standards of research excellence and integrity in architecture and philosophy. Former Assistant Director of the Dutch 3.TU Center for Ethics and Technology. Koller's research specializes in philosophy of technology, philosophy of architecture, and analytic philosophy of language.

Chapter 4
Technology as a Practical Art

Sven Ove Hansson

Abstract The notion of practical arts goes back to antiquity. It covers what we today call technology but also many other types of activities, such as farming, manual crafts, cooking, housekeeping, sport activities, artistic work, and medicine. Interesting discussions on the practical arts can be found in medieval texts on knowledge classification for instance by Hugh of Saint Victor and Robert Kilwardby, and also in Renaissance and Enlightenment literature. In this chapter it is argued that the philosophy of practical arts should be revitalized and that the philosophy of technology should have a major role when this is done. Many of the topics studied in the latter discipline have interesting extensions to practical arts in general. Some examples are the relationship between action knowledge and factual knowledge, the epistemological roles of actions, explanations in practical knowledge, and the role of functional terms and descriptions. *The Empirical Turn in the Philosophy of Technology* provides an excellent starting-point for widening the philosophy of technology to a general philosophy of the practical arts.

Keywords Technology • Practical arts • Mechanical arts • Empirical turn • Hugh of Saint Victor • Robert Kilwardby

4.1 Introduction

In the year 2000 I had a new job. My task was to introduce philosophy as a new discipline for research and teaching at the Royal Institute of Technology, Sweden's largest technological university. I realized that we needed to connect philosophy with the engineering disciplines and the subject matter of technology. I asked around and read a lot. Unfortunately much of the philosophy of technology that I encountered had little connection with current developments in technology or with the concerns of engineers and others working directly with technology. But I also found

---

S.O. Hansson (✉)
Royal Institute of Technology, Stockholm, Sweden
e-mail: soh@kth.se

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_4

quite a few philosophers whose investigations had a focus on the work of engineers, in much the same way as philosophers of science tend to focus on the work of scientists. Several of them were just finishing a new book, *The Empirical Turn in the Philosophy of Technology*. That book became my most important guide in developing philosophy at our university. And I am only one of many philosophers who have been influenced by it. Its impact has been immense in the philosophy of technology. Today, the discipline as a whole has much stronger connections with engineering research and with other philosophical disciplines, in particular the philosophy of science. I believe this is largely because so many of us have heeded the pleas in this book to investigate rather than speculate, and to focus on the actual technological practices that are readily available for our studies.

Obviously, a philosophy of technology is only possible if there is a concept of technology. Since that concept is about 200 years old, the philosophy of technology is a rather young discipline. However, it has a highly interesting prehistory. From ancient times to the early modern age philosophers discussed the nature and the social role of the "practical arts". This was a much wider concept that covers not only what we would today call "technology" but also many other human activities. In Sect. 4.2 I will provide a background on the practical arts and how they were discussed by philosophers. In Sect. 4.3 I will propose that the philosophical study of the practical arts should be resumed and that the philosophy of technology, in particular in the tradition from *The Empirical Turn*, provides us with most useful tools for doing so. Section 4.4 concludes.

4.2 The Practical Arts

4.2.1 Origins

The ancient world had no concept corresponding to what we today call technology. However, the ancients were aware that we need knowledge in order to achieve our goals in practical life. Activities based on such knowledge were classified by Aristotle as productive arts. He explained the concept as follows:

> Now since building is an art and is essentially a reasoned state of capacity to make, and there is neither any art that is not such a state nor any such state that is not an art, art is identical with a state of capacity to make, involving a true course of reasoning. All art is concerned with coming into being, i.e. with contriving and considering how something may come into being which is capable of either being or not being, and whose origin is in the maker and not in the thing made; for art is concerned neither with things that are, or come into being, by necessity, nor with things that do so in accordance with nature (since these have their origin in themselves). (Aristotle 1984, pp. 1799–1800)[1]

---

[1] *Nicomachean Ethics* VI:4, 1140a. Translation by W.D. Ross, revised by J.O. Urmson. The third word in the quotation was rendered "architecture" in Ross's original version.

In ancient Greece, the productive arts were considered inferior to other activities, and they were primarily performed by slaves and others with low status.[2] They were contrasted with the liberal arts, i.e. the arts related to reading, writing and arithmetic that were considered suitable for the education of free citizens. The Romans were equally negative towards practical activities, and they used various pejorative terms to denote them, such as *artes illiberales*, *artes vulgares*, *artes sordidae* and *artes banausicae* (Van Den Hoven 1996, pp. 90–91; Ovitt 1983; Tatarkiewicz 1963; Whitney 1990).

In the Middle Ages, the most common term for these arts was *artes mechanicae*, a phrase that seems to have been introduced by Johannes Scotus Eriugena (c.815–c.877) in his commentary on Martianus Capella's allegorical text on the liberal arts, *De nuptiis Philologiae et Mercurii* (On the Marriage of Philology and Mercury; Noble 1997). Today, this phrase is standardly translated as "mechanical arts", and it is often assumed to refer to what we today call technology. This interpretation may seem obvious, given the modern meaning of the word "mechanical". However, although the word has its origins in a Greek root that relates to machines, in the Middle Ages it acquired the meaning "concerned with manual work; of the nature of or relating to handicraft, craftsmanship, or artisanship". The old sense of "relating to machines" had disappeared by this period, and its later reappearance was probably the result of a learned reconstruction that took place in the sixteenth century (Oxford English Dictionary). Therefore *artes mechanicae* should preferably be translated as "manual arts" rather than "mechanical arts".

4.2.2 Hugh of Saint Victor and the Classificatory Tradition

When medieval authors mentioned the *artes mechanicae* the tone was usually condescending if not condemning. Like their ancient predecessors, they saw these activities as much less worthy than their own occupations. However, there were exceptions. Some medieval authors voiced a more positive appraisal of the *artes mechanicae*. The most famous example is the Saxon theologian Hugh of Saint Victor (c.1096–1141) whose *Didascalicon* contains a classification that divides the *artes mechanicae* into seven categories:

1. *Lanificium*: weaving, tailoring
2. *Armatura*: making of weapons, buildings, and metallic objects
3. *Navigatio*: trade on water and land
4. *Agricultura*: agriculture, horticulture, cooking

---

[2] In Aristotle this is most clearly expressed in *Politics*. He said that "in the state which is best governed and possesses men who are just absolutely… the citizens must not lead the life of artisans or tradesmen, for such a life is ignoble and inimical to excellence" (1328b36–39, cf. 1329a35–40). He admitted that "men must be able to engage in business and go to war, but leisure and peace are better; they must do what is necessary and indeed what is useful, but what is honourable is better" (1333b1–b3). (Aristotle 1984, pp. 2108–2109 and 2116. Translation by Benjamin Jowett.)

5. *Venatio*: hunting, food production
6. *Medicina*: medicine and pharmacy
7. *Theatrica*: theatre, music, gymnastics and games (Hugonis de S[ancto] Victore 1854, pp. 760–763. Cf. Hoppe 2011, pp. 40–41)

The reason why Hugh summarized all the practical arts under only seven headings was that he wanted to create a parallel with the liberal arts that were traditionally divided into seven categories (grammar, logic, rhetoric, arithmetic, geometry, music theory, and astronomy). Hugh emphasized that, just like the liberal arts, the mechanical ones could contribute to human wisdom and blessedness (Weisheipl 1965, p. 65).

Hugh's list confirms that the concept of *artes mechanicae* was much wider than our modern concept of technology. Only about half of the items on his list would be classified as technology today. Warfare, trade, hunting, medicine, games, and theatre playing are the clearest examples of items not so classified.

Most medieval discussions on the practical arts followed Hugh of Saint Victor in treating them in the context of a general classification of human knowledge. The typology of knowledge was a popular theme in the Middle Ages. A large number of classification schemes were presented in which the disciplines were organized in groups and subgroups, thus giving rise to a tree-like structure. Such *divisiones scientiarum* (*divisiones philosophiae*) served to identify areas that were deemed worthy of scholarly efforts and/or suitable for inclusion in educational curricula (Ovitt 1983; Dyer 2007).

Three terms were used interchangeably as umbrella terms for all knowledge: *scientia* (science), *philosophia* (philosophy), and *ars* (art). Etymologically one might expect a clear distinction between the three terms. *Scientia* is derived from *scire* (know), which was used primarily about knowledge of facts. *Philosophia* is of Greek origin and literally means "love of wisdom", but it was often interpreted as systematic knowledge and understanding in general, covering both empirical facts and more speculative topics such as existence and morality. *Ars* referred to skills, abilities, and craftsmanship. These three terms were all regularly used to cover all kinds of knowledge. Therefore, the practical arts (*artes mechanicae*) were described as included in all three concepts (Ovitt 1983; Freedman 1994; Covington 2005). It is perhaps particularly surprising to find that *philosophia* and its cognates in other languages were taken to encompass all kinds of knowledge, including practical craftsmanship. This usage of the English word "philosophy" was common as late as in the eighteenth century (Tonelli 1975; Freedman 1994).

4.2.3 Robert Kilwardby

The English philosopher and theologian Robert Kilwardby (c.1215–1279) was another major contributor to the medieval discussion on the nature of human knowledge and the classification of the knowledge disciplines. He was much inspired by

Hugh of Saint Victor, not least Hugh's discussion of the practical arts (Maierù 2013). In his widely read *De Ortu Scientiarum* (On the origin of the sciences), Kilwardby distinguished between two major types of philosophy.[3] One of these was *philosophia speculativa*, which contained metaphysics, mathematics, and natural philosophy (physics). The word *speculativus* should not be interpreted in the modern sense of hypothetical or groundless; it derives from the Latin *speculor* which means to observe, examine or explore. The best translation is probably "theoretical".

His other major category was philosophy on human matters (*de rebus humanis*) that was further subdivided into two subcategories. One of these was the study of language and the other was *philosophia activa vel practica* (active or practical philosophy). The latter in its turn had two subdivisions, namely *philosophia mechanica* (manual philosophy, or the practical arts) and *philosophia ethica* (ethical philosophy) (Kilwardby 1976; Maierù 2013, pp. 359 and 379; Sirridge 1988).

Kilwardby gave a fairly detailed account of Hugh of Saint Victor's classification of the mechanical arts. He was positive to most parts of Hugh's typology, but there was one part that he could not accept. One of Hugh's seven classes was the theatric arts (*theatrica*). In Kilwardby's view, such activities had no place on a list of the (useful) practical arts. Theatre, he said "does not seem to me to be something to be presented to Catholics, but rather to be detested and fought against" (*non videtur mihi ponenda apud catholicos*, *sed magis detestanda et impugnanda*) (Kilwardby 1976, p. 131). (By a catholic was at that time meant an adherent of the Roman Church in contradistinction to the Eastern or Orthodox Churches.) Therefore he removed theatrics from the list of seven arts. The parts of it that he could accept, namely "that which is allowed for Catholics to play, such as the lyre, the trumpet, the flute and the like" (Kilwardby 1976, p. 132) was instead placed under the rubric of medicine. To ensure that the practical arts were still seven in number he elevated architecture so that it became one of the seven main categories. However, at the same time he pointed out that it was not really necessary for the practical arts to be exactly seven like the liberal arts (Kilwardby 1976, p. 133).

Kilwardby took it for self-evident that the practical arts belonged to philosophy, but he considered them to be science (*scientia*) only in part. This was because he considered science to consist only of that which could be known with certainty. With "true and certain science" (Kilwardby 1976, p. 137) he meant a discipline that is "based on propositions which are necessary per se and in themselves" and on the conclusions drawn from them (Kilwardby 1976, p. 135).[4] Therefore knowledge that was science to the highest degree could only be found in metaphysics and mathematics. He assigned this status to metaphysics due to the "dignity of the subject" and to mathematics for an arguably more convincing reason, namely the "certainty of the mode of demonstration" (Kilwardby 1976, p. 137). Physics was less of a

---

[3] This division of philosophy into two types can also be found in Boethius. (Maierù 2013, p. 353)

[4] Here he follows Eustratius' comment on Aristotle's *Nicomachean Ethics*. See the editorial footnote in Kilwardby (1976, p. 135).

science than these two, and ethics even less so than physics. The crafts were sciences to the least degree.

However, he had a rather sophisticated view of the relationship between theoretical and practical knowledge that went beyond these simple classifications. In a highly interesting passage he emphasized that the two types of knowledge depend on each other:

> Having said something separately concerning the speculative part of philosophy and something about the practical part, now it is suitable to say something about them in comparison with each other. I ask therefore in what way they are distinguished according to their degree of speculative philosophy and praxis, since those which are practical are also speculative – one should namely first consider by speculative virtue what one ought to perform in practical virtue – and, conversely, the speculative sciences are not without praxis. Does not, in fact, arithmetic teach how to add numbers to each other and to subtract them from each other, to multiply and divide and draw out their square roots, all of which are operations? Similarly, does not music teach to play the lute and flute and things of this sort? Again does not geometry teach how to measure every dimension, through which both carpenters and masons work? Again, does one not know the time for navigation and sowing and things of this sort through astronomy? It can thus be seen that every single science said to be speculative is also practical. Therefore it is clear that the speculative sciences are practical and the practical speculative. (Kilwardby 1976, p. 138)[5]

The relationship between theoretical and practical knowledge was also discussed by the Italian philosopher Jacopo Zabarella (1533–1589) in his *Opera Logica* (1578). However, he reached a different conclusion than Kilwardby. In Zabarella's view, the practical arts can learn from the theoretical (or as he said "contemplative") ones, but not the other way around. He also denied that the practical arts are at all sciences (*scientiae*). Interestingly he used medicine as an example of this. Since medicine serves a practical end, namely health, rather than knowledge for its own sake, it cannot in his view be a science. If studies of the human body and its diseases were performed with the sole purpose of knowing more about them, rather than as a means to cure diseases (what we would today call human biology), then that would be a science, but it would not be medicine any more (Mikkeli 1997, pp. 221–222).

4.2.4 Increasing Appreciation of the Practical Arts

In the early modern age, several authors expressed a much higher appreciation of the practical arts than what had been common previously. In the utopian tract *Civitas Solis* (City of the Sun, 1623) the Italian friar Tommaso Campanella (1568–1639) declared that manual work was as valuable and dignified as any other work. In contrast, he saw the idleness of the nobility as both undignified and vicious. In James Harrington's (1611–1677) *The Commonwealth of Oceana* (1656) the usefulness of education in the practical arts was emphasized. In the schools of Oceana, the

---

[5] The translation follows Whitney (1990, p. 120) but a few changes have been made.

students were taught "mechanicks" by which was meant farming, manufacturing, and merchandizing. But probably the most important step towards scholarly recognition of the practical arts was taken in the great French *Encyclopédie*, published from 1751 to 1772, that was the most influential literary output of the Enlightenment. One of its achievements was the incorporation of the practical arts into the edifice of learning. Its volumes contained detailed descriptions of the work processes of many practical trades and a large number of engravings showing workshops, tools and work processes. In his introduction to the *Encyclopédie*, the mathematician Jean Le Rond d'Alembert (1717–1783) claimed that the mechanical arts were no less worthy pursuits than the liberal arts.

> The mechanical arts, which are dependent upon manual operation and are subjugated (if I may be permitted this term) to a sort of routine, have been left to those among men whom prejudices have placed in the lowest class. Poverty has forced these men to turn to such work more often than taste and genius have attracted them to it. Subsequently it became a reason for holding them in contempt – so much does poverty harm everything that accompanies it. With regard to the free operations of the mind, they have been apportioned to those who have believed themselves most favoured by Nature in this respect. However, the advantage that the liberal arts have over the mechanical arts, because of their demands upon the intellect and because of the difficulty of excelling in them, is sufficiently counterbalanced by the quite superior usefulness which the latter for the most part have for us. It is this very utility which has reduced them forcibly to purely mechanical operations, so that the practice of them may be made easier for a large number of men. But society, while rightly respecting the great geniuses which enlighten it, should in no wise debase the hands which serve it. (d'Alembert 1751, p. xiij)

4.2.5 The Fine Arts

In the modern era two important subcategories were separated out from the practical arts. The first of these was the concept of fine arts (*beaux arts*), today often just called "art" or "the arts". Obviously, sculpture, painting, music, dance, drama, and literature all go back to prehistoric ages. However, they have not always been seen as comprising a special kind of human endeavour distinct from practical, intellectual, or religious activities. For instance, no division was usually made between that part of a potter's work that consists in making the pottery durable and fit for use and that part which consists in making it appealing to the eye. The occupations that we today call artistic, such as sculpture, painting, and music, were usually treated in the same way as other qualified manual trades, not least concerning the social status of those performing the work.

In ancient Greece, reference was sometimes made to some of the arts as being "imitating". For instance, in *Poetics* Aristotle noted that "[e]pic poetry and tragedy, as also comedy, dithyrambic poetry, and most flute-playing and lyre-playing, are all, viewed as a whole, modes of imitation" (Aristotle 1984, p. 2316, 1447a14–15). In the same context he mentioned "the dancer's imitations" (ibid, 1447a27) and the arts that "imitate and portray" by means of colour and form (ibid, 1447a19–20).

This passage strongly indicates that there was a concept of "imitative arts" in the ancient world (Young 2015) However, although they had the concept of an imitative art, this was not a category used to classify the various arts.[6]

The ancient concept of imitative arts does not seem to have had much influence on medieval writers. For instance, in the medieval knowledge classifications poetry was grouped with grammar and rhetoric, the visual arts with the other manual crafts, and music with mathematics (due to the mathematical nature of music theory; cf. James 1995). None of the knowledge classifiers seem to have considered forming a group consisting of what we today call the artistic disciplines.

The concept of imitative arts was revived in the sixteenth century, sometimes under the name "imitative", sometimes under names such as "the better arts"[7] or "les beaux arts" (the fine arts) (Young 2015). The latter term seems to have been invented by Toussaint Rémond de Saint-Mard (1682–1757) (Rémond de Saint-Mard 1734, p. 314). It was popularized by Charles Batteux (1713–1780), professor of philosophy in Paris, in his influential book from 1746, *Les beaux arts réduits à un même principe* (The Fine Arts Reduced to a Single Principle). Batteux has been credited with being the first to group the fine arts together and separating them clearly from the other practical arts (Kristeller 1980). The single principle that he referred to in the title of his book was the imitation of nature, the same principle that already Aristotle had referred to as constitutive of these arts (Batteux 1746, pp. 5–7).

> Thus we can divide the arts into three sorts with regard to the purpose proposed.
>
> For some, their purpose is the needs of man, whom nature seems to abandon as soon as he is born, exposed to cold, hunger and a thousand difficulties, insisting that remedies and protection be the price of our invention and work. Thus the mechanical arts were born.
>
> The object of the second group is pleasure. They were born only in the womb of joy and of feelings that plenty and tranquility produce: these are called the fine arts par excellence. Such are music, poetry, painting, and the art of gesture or dance.
>
> The third category includes the arts that are both useful and agreeable: these include eloquence and architecture. Need brought them into being, taste perfected them; they occupy a middle position between the two other types… (Batteux 1746, pp. 5–7)

Batteux's concept of the "fine arts" was well received by his contemporaries. The authors of the *Encyclopédie* contributed much to its general acceptance. In his introduction, d'Alembert made extensive use of Batteux's terminology. According to d'Alembert, a major difference between the fine arts and more theoretical learning was that the former were much less rule-bound, and therefore much more dependent on individual inspiration or genius.

> Among the liberal arts that have been reduced to principles, those that undertake the imitation of Nature have been called the Fine Arts because they have pleasure for their principal

---

[6] In the *Metaphysics* Aristotle distinguished between those arts that were "directed to the necessities of life" and those that were devoted to "its recreation". However, the only example that he mentioned of the latter was "the mathematical arts", which shows that this category is different from his notion of an imitative art and also different from the modern concept of the fine arts. (981b27–24. Aristotle 1984, p. 1553. Translation by W.D. Ross.)

[7] This is Giorgio Vasari's phrase ("le migliori arti"). It was somewhat freely translated by Mrs. Foster as "the higher arts" (Vasari [1550] 1855, p. 275).

> object. But that is not the only characteristic distinguishing them from the more necessary or more useful liberal arts, such as Grammar, Logic, and Ethics. The latter have fixed and settled rules which any man can transmit to another, whereas the practice of the Fine Arts consists principally in an invention which takes its laws almost exclusively from genius. The rules which have been written concerning these arts are, properly speaking, only the mechanical part. Their effect is somewhat like that of the telescope; they only aid those who see. (d'Alembert 1751, p. xiij.)

4.2.6 Technology

The other new subcategory that was distinguished among the practical arts was technology. The word "technology" is of Greek origin, and originally meant knowledge about the arts (The Greek word *techne*, τέχνη, was very close in meaning to the Latin *ars*.). The word was apparently introduced into Latin by Cicero (Steele 1900, p. 389; Cicero [c. 45 B.C.] 1999, *Epistulae ad Atticum* 4:16). However, it does not seem to have been much used until Peter Ramus (1515–1572) began to use it in the sense of knowledge about the relations among all the arts. It received its modern meaning through two major changes in meaning. The first of these changes was an increased emphasis in its use on the skills and devices of craftspeople. This was a gradual change that received a clear expression in 1829 when the American physician and scientist Jacob Bigelow (1787–1879) published his *Elements of Technology*, in which he defined technology as "the principles, processes, and nomenclatures of the more conspicuous arts, particularly those which involve applications of science" (Tulley 2008; Sebestik 1983). This sense became increasingly dominant, and in 1909 *Webster's Second New International Dictionary* defined technology as "the science or systematic knowledge of industrial arts, especially of the more important manufactures, as spinning, weaving, metallurgy, etc" (Tulley 2008). By then technology had acquired a more limited sense. It referred to knowledge about that which is done with tools and machines.

The more precise demarcation that the word "technology" acquired in this process seems to have been influenced by the delimitation of the new engineering educations that emerged in Europe in the nineteenth century. Previously "engineer" had been a military title.[8] But in the nineteenth century schools for the education of civil engineers were initiated throughout Europe, for instance the Technological Institute in Stockholm, founded in 1827 (now the Royal Institute of Technology, the institution where this text is written). Originally, these were schools for young craftsmen in the towns. Therefore, their education had its focus on the tools, machines and work processes employed by this class of people. For the most part this excluded the

---

[8] "Engineer" derives from the latin *ingenium* that was used in the classical period for a person's talent or inventiveness, but could also refer to a clever device or construction. In the Middle Ages, *ingenium* was a general term for catapults and other war machines for sieges. A constructor or master builder of such devices was called *ingeniarius* or *ingeniator*. (Bachrach 2006; Langins 2004)

tools, machines, and processes that were used by farmers and farm workers, women, and members of the "higher" professions such as pharmacists and surgeons. The usage of the word "technology" followed the same pattern. We still do not consider farming, pharmacy, cooking, or surgery as technological occupations, although they involve equally extensive and sophisticated use of tools and machines as many of the occupations so classified. This means that the demarcation of technology is arbitrary in the sense of being much influenced by historical and social contingencies.

The second change in the meaning of "technology" took place primarily in the English language in the first half of the twentieth century. Increasingly often, "technology" referred to the actual tools, machines, and procedures used to produce material things, rather than to knowledge about them. The earliest example of this meaning given in the *Oxford English Dictionary* is a text from 1898 about the coal-oil industry, according to which "a number of patents were granted for improvements in this technology, mainly for improved methods of distillation" (Peckham 1898, p. 119). Today this is the dominant meaning of the word in English. As Joost Mertens noted, "[i]n English usage, 'technology' normally refers to instrumental practices or their rules and only exceptionally to the scientific description, explication or explanation of these practices." (Mertens 2002) The new sense of the word has spread to other languages. For instance, in French, German, Dutch, and Swedish it is commonly used both in the old sense of knowledge about tools, machines and processes and in the new sense, taken from English, of the tools, machines and processes themselves.[9] According to the *Svenska Akademiens Ordbo*k, the Swedish counterpart of the *Oxford English Dictionary*, this usage became common in Swedish in the 1960s.

In more recent years, the meaning of "technology" has evolved in a way that seems to have followed the development of curricula in engineering schools. Two major extensions of its meaning took place in the second half of the twentieth century. First, with the development of computer and information technology, a wide range of programming and other software-related activities became recognized as technological. Secondly, through the equally rapid development of biotechnology, many activities based on biological knowledge are now considered technological, and new types of artefacts, namely various forms of modified biological organisms, are regarded as technological products. However, the arbitrariness referred to above still persists: There are still areas of knowledge, such as farming and surgery, that we do not describe as technological although they have as much focus in the use of tools and machines as most of the areas that we call technological.

---

[9] The French, German, Dutch, and Swedish languages all have a shorter word (technique, Technik, techniek, teknik) that refers to the actual tools, machines and practices themselves.

4.3 A New Philosophy of the Practical Arts

The notion of practical arts is quite wide. It includes the two subareas that broke away from it, the fine arts and technology. It also includes many other types of practical knowledge, for instance farming, medicine, sport activities, car driving, cooking, and housekeeping. In what follows I hope to show that they have so many interesting philosophical issues in common that we have good reasons to revive the philosophical study of the practical arts as a general category.[10]

A revival of the philosophy of practical arts can have several starting-points, of which I would like to emphasize three. First, some of the medieval discussions (such as Kilwardby's discussion of the relation between practical arts and theoretical knowledge) provide useful beginnings for new developments. Secondly, many topics from the philosophy of technology, especially after its empirical turn, can be generalized so that they refer to the practical arts in general rather than only to technology. Thirdly, philosophical studies of the fine arts (usually labelled aesthetics or philosophy of art) also provide opportunities for generalizations to the other practical arts.[11]

There is an interesting difference between the philosophy of technology and that of the fine arts. The latter has traditionally an almost exclusive focus on the perspective of individual consumers (recipients) of (artistic) products.[12] In fact, the term "aesthetics" derives from a Greek word meaning to perceive or to sense. In contrast, the philosophy of technology has usually been dominated by other perspectives, such as social consequences and (following the empirical turn) the perspectives of producers or inventors. These various approaches are complementary, and I believe it would be useful to have many more studies of the production of fine arts (making music, writing a novel, painting etc.) and also of the individual reception and use of technology.

In what follows I will propose seven topics for a generalized philosophy of the practical arts. These topics are all tentative. The list has a certain preponderance of themes that are close to philosophy of technology in the empirical turn tradition.

---

[10] In a somewhat similar vein the German historian of technology Otto Mayr proposed that research should be conducted on "historical interactions and interchanges between what can roughly be labelled 'theoretical' and 'practical' activities, that is, between man's investigations of the laws of nature and his actions and constructions aimed at solving life's material problems" (Mayr 1976, p. 669).

[11] It should be observed that neither the (fine) arts nor technology are well-defined concepts today. Current developments for instance in conceptual, relational, and virtual art contribute to making the concept of "fine arts" problematic. At the same time, developments in technology such as autonomous software agents, biological engineering, and artificial life overstep the limits of what we have previously meant by technology. The definitional lability of the traditional concepts can be seen as an additional reason to try out other ways to categorize the phenomena that they are supposed to cover.

[12] There are a few exceptions, such as Bengt Edlund's analyses of the aesthetics of playing (as opposed to listening to) instrumental music. Bodily (proprioceptive) perception can give rise to aesthetic experiences for the musician that are inaccessible to the listener (Edlund 1996a, b, 2003).

4.3.1 The Relationship Between Action Knowledge and Factual Knowledge

The distinction between basic and applied science is commonly used, but its meaning is far from clear. The term "applied science" can be understood in two rather different ways that are often confused. In one sense, applied science denotes the use of available science for practical purposes. In that sense, applied science is essentially uncreative and does not lead to the generation of new scientific knowledge. This interpretation of the term has been at focus in most of the comments emphasizing that technology is not applied science (Bunge 1988; Lelas 1993; Mitcham and Schatzberg 2009). However, there is also another sense of "applied science", namely science devoted to applied or practical problems. In this latter sense, applied science can be creative and produce genuinely new knowledge. This usage of the term also has a fairly long history of usage (Kline 1995; Gooday 2012). (Arguably, it would nevertheless be clearer to refer to "practice-guiding" or "action-guiding" science rather than to use the term "applied" in this sense as well.) We can think of basic science as science providing us with knowledge about what the world is like, and applied (action-guiding) science as science telling us how we can change the world to achieve various practical goals.

Basic and applied science are commonly taken to be mutually exclusive categories. However, much scientific research has both the quality of providing us with new understanding and that of providing new knowledge of how we can achieve practical results (Stokes 1997). For instance, one and the same scientific study may be helpful both for understanding how the human body works and for developing treatments against some disease. Similarly, one and the same study can improve our understanding of interactions among living organisms and our knowledge of how to prevent environmental damage.[13] Therefore it may be better to classify scientific knowledge according to the two criteria of (a) its contribution to our understanding of the world and (b) its contribution to our knowledge of how we can change the world. As already indicated, one and the same scientific discovery or invention can satisfy both these criteria to a high degree.

This approach can be extended to cover all forms of knowledge, not only scientific knowledge. Just like their scientific counterparts, non-scientific factual knowledge and non-scientific action knowledge have considerable overlaps. For instance, I know that a certain musical symbol denotes a semiquaver rest. This is factual knowledge, but it is also action-guiding when I play music. The classification of such knowledge is far from self-evident. Should we describe it as factual knowledge that can be used to produce action knowledge, or should it be classified as both

---

[13] The practical usefulness of knowledge changes over time. Symbolic logic was once without application; today it is the basis of computer technology (MacLane 1965). Number theory, once the epitome of pure mathematics, is now applied in cryptography (Schroeder 2009). In recent years, research on photosynthesis has become practically relevant through efforts to develop artificial photosynthesis (Herrero et al. 2011) and to improve the photosynthetic efficiency of agricultural plants (Ku et al. 2001).

factual knowledge and action knowledge? Depending on how we perform this classification, the relationships between factual knowledge and action knowledge may come out differently, in particular in terms of how knowledge of one type might depend on the other. This can be seen as an extension of the issue how science and technology relate to each other, and insights gained in the philosophy of technology should be useful inputs into its investigation.

4.3.2 The Specialization and Systematization of Knowledge

Science is the result of a systematization of knowledge and knowledge-conducive activities. The distinction between scientific and non-scientific knowledge does not (or at least should not) refer to the subject matter but to other features such as systematicity, revisability, and the encouragement of criticism (Hansson 2013b).

Knowledge and beliefs can be systematized in other forms than science; for instance, hunters in some societies that lack a written language have developed highly specialized and strictly fact-oriented ways to express and discuss their knowledge about the behaviour of prey (Blurton Jones and Konner 1976). Both factual knowledge and action knowledge can be systematized in various ways, only some of which are described as science. There are many types of systematized knowledge in the practical arts, for instance the instructions given to someone learning a practical craft, a musical instrument, a conjuring trick, or a sport. Such systematizations have some of the features of a science, but usually not all of them. (The exceptions are those practical arts, such as clinical medicine and some forms of technology, that have become so imbued with science that they can be described as science-based.) The characteristics of various types of knowledge systematizations should be studied in order to provide us with a more general understanding of the nature of collective knowledge, and also to better understand the relationship between science and other collectively organized repositories of knowledge.

4.3.3 The Epistemological Role of Actions

Classically, we refer to two major sources of human knowledge: our sense organs and our reasoning capacity. These two sources of knowledge are related to the two major ideas about the foundations of scientific knowledge: empiricism that puts emphasis on what we learn from our sense organs and rationalism that puts emphasis on what we learn by rational thinking. However, there is also a third source of knowledge: our actions. The importance of action for action knowledge is obvious, but it is not difficult to see that it is essential for factual knowledge as well. We can use the term "epistemic action" as a general term for actions performed with the intention to obtain knowledge. The classification and analysis of knowledge-conducive actions remains to be developed, but three types of such action are obvious candidates, namely actions of trying, exercising, and investigating.

In order to find out how to do something, we *try* different ways to do it. We do this in all kinds of everyday situations. When acts of trying are performed more systematically, for instance in agricultural field trials or in the clinical trials of medicine, they take the form of directly action-guiding experiments (Hansson 2015).

Exercising differs from trying in that it consists in doing something that one already has some minimal ability to do. By doing it repeatedly with close attention to the result we learn to do it better or more effortlessly. Exercising is an important component in the acquisition of tacit knowledge (Hansson 2013a).

Investigating actions are actions we perform to support or facilitate the use of our sense organs. We perform investigations to achieve both factual knowledge and action knowledge. There are several subclasses of investigating actions. We uncover and move objects to make them available for observation (for instance in archaeological excavations and biological field excursions). We manipulate objects to make them observable (for instance by slicing and staining objects for microscopy), and we organize the conditions of our observations in order to make them suitable for conclusions about regularities in nature (i.e., we construct experiments). We also build instruments for our observations (Boon 2015).

All these types of epistemic actions are used to obtain both factual knowledge and action knowledge, both in science and in various practical arts. Although some types of epistemic actions have been discussed rather extensively we still lack a general account of epistemic actions, how they are used to obtain different types of knowledge and how they relate to the other two major sources of knowledge, our senses and our reasoning abilities. Recent studies in the epistemology of technology provide an excellent starting-point for such investigations.

4.3.4 The Role of Explanation in the Practical Arts

In scientific work, explanations are expected to be compatible with the current state of knowledge; otherwise they are renounced. Per Norström (2013) has shown that this is not true in engineering where explanations are often retained even if they are known to be based on incorrect ideas. The idea that vacuum sucks is a clear example of that. This way of talking is common not only in everyday discussions among engineers but also in patent applications and official documents. Similarly, heat is often described as a substance that can move between objects, and in discussions about building isolation, cold is similarly described as something that can “leak” into a building. In his study of such explanations, Norström pointed out that due to the practical purposes of engineering, there may be nothing wrong in using scientifically untenable explanations as long as they fill their purposes.

> Technology is about creating artefacts and solving problems, while science is primarily about describing and explaining phenomena in the world… Technological theories do not become obsolete because their foundations have been falsified. They become obsolete

> when nobody has use for them any longer. Among the theories and models used by practicing engineers and technicians today there are many that are based on obsolete science. Thereby they lack scientific justification, but they are nonetheless useful. (Norström 2013, p. 378)

Most other practical arts have less developed connections with science than what technology has. Therefore an investigation of explanations in some of these practical arts might very well show an even greater prevalence of non-scientific explanations than in technology. Whereas there is a substantial literature on explanations in science much less is known about explanation in the practical arts.

4.3.5 Activating the Audience

One common trend in the arts is the activation of the audience. Theatre audiences are invited to take part in the performance, television viewers take part in programs for instance by voting, and newspaper readers write comments on the article. Experiments are made in interactive literary forms such as hypertext fiction. New, highly interactive art forms are being developed, based on virtual reality and computer gaming (Charles et al. 2002; Cover 2006; Morales-Manzanares et al. 2001; Paradiso et al. 1999; Rowe 1999).

At the same time, consumer co-creation increases in many branches of commerce and technology. By this is meant that consumers create data, such as book reviews on a bookshop website, surf patterns on search machines, and various forms of data that are produced in online playing of computer games. This information is used for free by the company that owns the website to improve their commercial services. Consumer co-creation can be seen as a new, more active and more rewarding way to use the services in question. But it can also be seen as unpaid work for commercial enterprises, with little or no information on how the work output will be used (Banks and Deuze 2009; Banks and Humphreys 2008; Flowers 2008).

These two forms of interaction, audience participation and consumer co-creation, have much in common but they have mostly been discussed as separate issues. A more unified approach can hopefully lead to a better understanding of the phenomenon, including its ethical aspects.

4.3.6 Functional Terms

One of the major modern achievements in the philosophy of technology is the analysis of technological function. *The Empirical Turn* provides many examples of how useful a focus on functional terms and descriptions can be in philosophical investigations of technology. This analysis has been further developed and many new insights have been reported. One major example is Peter Kroes's careful distinction between functional kind, ascribed function, and functionality (Kroes 2012).

Interestingly, functional terms are also quite common in some of the other practical arts. In medicine, therapies are largely categorized according to their function (analgesics, antibiotics, bronchodilators etc.). In musical theory, chords are classified according to their "harmonic function", in sports team members are classified according to their functions in the team (goalkeeper, forward), farmers classify animals according to their function (dairy cattle, beef cattle, workhorse), cooks classify ingredients according to their function (spice, preservative, sugar substitute) etc. These uses of functional terminology are quite different in nature. It is an open question how much they have in common with each other and with the uses of functional terms in technology. However, this is a question that would be well worth investigating. The research on technological function whose early achievements are reported in *The Empirical Turn* is the obvious starting-point for such investigations.

4.3.7 Philosophy of Ends

Traditionally, decision analysis has been based on the assumption that goals are inputs that should just be taken for given, and the analysis should focus on how the goals are best achieved. This is also the way in which design criteria in engineering design are usually conceived. However, an approach that takes practical goals to be incontestable is much too crude. Most goal-driven practical activities can gain a lot from a critical discussion of current or proposed goals, for instance in terms of how conducive the setting of a goal is to its attainment. There are various reasons why a goal can fail to achieve its intended purpose, and some of these reasons may be sufficient to justify that we change or give up the goal (Edvardsson and Hansson 2005). However, largely due to the common oversimplification just referred to, the philosophy of goal-setting rationality has as yet not attracted much attention among researchers, and many basic issues remain to investigate.

The practical arts are by definition devoted to the attainment of practical goals. Therefore, goal-setting and its rationality is a most suitable topic for studies in a re-established philosophy of the practical arts.

4.4 Conclusion

The concept of technology is only about 200 years old, and its delimitation is marked by historical contingencies. I have proposed that we should have a close look at the older and much wider concept of the practical arts. Many of the topics studied in the philosophy of technology have interesting extensions that refer to practical arts in general. Some examples are the relationship between action knowledge and factual knowledge, the collective systematization of knowledge, the epistemological role of actions, explanations in practical knowledge, and the use of functional terms and descriptions. The Empirical Turn can be applied to a wider category of human activities than those covered by the term "technology". For this we are well prepared, thanks to the remarkable advances in the philosophy of technology in the last two decades.

Acknowledgement I would like to thank Maarten Franssen and Peter Kroes for unusually valuable comments on an earlier version of this text.

References

Aristotle. (1984). In J. Barnes (Ed.), *Complete works*. Princeton: Princeton University Press.

Bachrach, D. S. (2006). English artillery 1189–1307: The implications of terminology. *English Historical Review, 121*(494), 1408–1430.

Banks, J.A., & Deuze, M., Eds. (2009). Co-creative labour. Special issue of the *International Journal of Cultural Studies*, *12*(5), 419–547.

Banks, J. A., & Humphreys, S. M. (2008). The labour of user co-creation. Emerging social network markets? *Convergence: The International Journal of Research into New Media Technologies, 14*, 401–418.

Batteux, C. (1746). *Les Beaux Arts Réduits à un Même Principe*. Paris: Durand.

Blurton Jones, N., & Konner, M. J. (1976). !Kung knowledge of animal behaviour. In R. B. Lee & I. DeVore (Eds.), *Kalahari hunter-gatherers* (pp. 326–348). Cambridge, MA: Harvard University Press.

Boon, M. (2015). The scientific use of technological instruments. In S. O. Hansson (Ed.), *The role of technology in science: Philosophical perspectives* (pp. 55–79). Dordrecht: Springer.

Bunge, M. (1988). The nature of applied science and technology. In V. Cauchy (Ed.), *Philosophy and culture: Proceedings of the XVIIth congress of philosophy* (Vol. II, pp. 599–604). Montréal: Éd. Montmorency.

Charles, F., Mead, S. J., & Cavazza, M. (2002). From computer games to interactive stories: Interactive storytelling. *Electronic Library, 20*, 103–112.

Cicero, M.T. (1999). *Epistulae ad Atticum*, vols. 1–4 (edited and translated by D.R. Shackleton Bailey). Loeb Classical Library. London: Harvard University Press.

Cover, R. (2006). Audience inter/active: Interactive media, narrative control and reconceiving audience history. *New Media and Society, 8*, 139–158.

Covington, M. A. (2005). Scientia sermocinalis: Grammar in medieval classifications of the sciences. In N. McLelland & A. Linn (Eds.), *Flores grammaticae: Essays in memory of Vivien Law* (pp. 49–54). Münster: Nodus Publikationen.

d'Alembert, J. l. R. (1751). Discours préliminaire des editeurs. *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers, 1*, 1751 (pp. i–xlv).

Dyer, J. (2007). The place of musica in medieval classifications of knowledge. *Journal of Musicology, 24*, 3–71.

Edlund, B. (1996a). On scores and works of music: Interpretation and identity. *British Journal of Aesthetics, 36*(4), 367–380.

Edlund, B. (1996b). Structural symmetry and proprioceptive patterns in music. *Symmetry: Culture and Science, 7*(2), 139–151.

Edlund, B. (2003). The phenomenology of fingering: Structure and ontology in Chopin's 'F-minor Etude' from 'méthode des Méthodes'. In I. Poniatowska (Ed.), *Chopin and his work in the context of culture* (Vol. 2, pp. 88–105). Warsaw: Polska Akademia Chopinowska.

Edvardsson, K., & Hansson, S. O. (2005). When is a goal rational? *Social Choice and Welfare, 24*, 343–361.

Flowers, S. (2008). Harnessing the hackers: The emergence and exploitation of outlaw innovation. *Research Policy, 37*, 177–193.

Freedman, J. S. (1994). Classifications of philosophy, the sciences, and the arts in sixteenth- and seventeenth-century Europe. *Modern Schoolman, 72*, 37–65.

Gooday, G. (2012). 'Vague and artificial': the historically elusive distinction between pure and applied science. *Isis, 103*, 546–554.

Hansson, S. O. (2013a). What is technological knowledge? In B. Skogh & M. J. de Vries (Eds.), *Technology teachers as researchers* (pp. 17–31). Rotterdam: Sense Publishers.

Hansson, S. O. (2013b). Defining pseudoscience – and science. In M. Pigliucci & M. Boudry (Eds.), *The philosophy of pseudoscience* (pp. 61–77). Chicago: Chicago University Press.

Hansson, S. O. (2015). Experiments before science: What science learned from technological experiments. In S. O. Hansson (Ed.), *The role of technology in science. Philosophical perspectives* (pp. 81–110). Dordrecht: Springer.

Herrero, C., Quaranta, A., Leibl, W., Rutherford, A. W., & Aukauloo, A. (2011). Artificial photosynthetic systems: Using light and water to provide electrons and protons for the synthesis of a fuel. *Energy and Environmental Science, 4*, 2353–2365.

Hoppe, B. (2011). The latin '*artes*' and the origin of modern 'arts'. In M. Burguete & L. Lam (Eds.), *Arts: A science matter* (Vol. 2, pp. 35–68). Singapore: World Scientific.

Hugonis de S[ancto] Victore (1854). *Opera Omnia, Tomus Secundus*. Edited by J. P. Migne. Lutetiae Parisiorum [Paris].

James, J. (1995). *The music of the spheres*. London: Abacus.

Kilwardby, R. (1976). In A. G. Judy (Ed.), *De Ortu Scientiarum*. Oxford: Oxford University Press.

Kline, R. (1995). Construing 'technology' as 'applied science': Public rhetoric of scientists and engineers in the United States, 1880–1945. *Isis, 86*, 194–221.

Kristeller, P. O. (1980). *Renaissance thought and the arts: Collected essays*. Princeton: Princeton University Press.

Kroes, P. (2012). *Technical artefacts: Creations of mind and matter. A philosophy of engineering design*. Dordrecht: Springer.

Ku, M. S. B., Cho, D., Li, X., Jiao, D.-M., Pinto, M., Miyao, M., & Matsuoka, M. (2001). Introduction of genes encoding C4 photosynthesis enzymes into rice plants: Physiological consequences. In J. A. Goode & D. Chadwick (Eds.), *Rice biotechnology: Improving yield, stress tolerance and grain quality* (Novartis Foundation Symposium, Vol. 236, pp. 100–116). New York: Wiley.

Langins, J. (2004). *Conserving the enlightenment: French military engineering from Vauban to the revolution*. Cambridge, MA: MIT Press.

Lelas, S. (1993). Science as technology. *British Journal for the Philosophy of Science, 44*, 423–442.

MacLane, S. (1965). Quality in science? *Bulletin of the Atomic Scientists, 21*, 7–11.

Maierù, A. (2013). Robert Kilwardby on the division of the sciences. In H. Lagerlund & P. Thom (Eds.), *A companion to the philosophy of Robert Kilwardby* (pp. 353–398). Leiden: Brill.

Mayr, O. (1976). The science-technology relationship as a historiographic problem. *Technology and Culture, 17*, 663–673.

Mertens, J. (2002). Technology as the science of the industrial arts: Louis-Sébastien Lenormand (1757–1837) and the popularization of technology. *History and Technology, 18*, 203–231.

Mikkeli, H. (1997). The foundation of an autonomous natural philosophy: Zabarella on the classification of arts and sciences. In D. A. Di Liscia, E. Kessler, & C. Methuen (Eds.), *Method and order in renaissance philosophy of nature: The Aristotle commentary tradition* (pp. 211–228). Aldershot: Ashgate.

Mitcham, C., & Schatzberg, E. (2009). Defining technology and the engineering sciences. In A. Meijers (Ed.), *Handbook of the philosophy of science* (Philosophy of technology and engineering sciences, Vol. 9, pp. 27–63). Amsterdam: Elsevier.

Morales-Manzanares, R., Morales, E. F., & Dannenberg, R. (2001). SICIB: An interactive music composition system using body movements. *Computer Music Journal, 25*(2), 25–36.

Noble, D. F. (1997). *The religion of technology: The divinity of man and the spirit of invention*. New York: Alfred A. Knopf.

Norström, P. (2013). Engineers' non-scientific models in technology education. *International Journal of Technology and Design Education, 23*, 377–390.

Ovitt, G., Jr. (1983). The status of the mechanical arts in medieval classifications of learning. *Viator, 14*, 89–105.

Oxford English Dictionary, online, http://www.oed.com.

Paradiso, J., Hsiao, K.Y., Hu, E. (1999). Interactive music for instrumented dancing shoes. In *Proceedings of the 1999 International Computer Music Conference* (pp. 435–456).

Peckham, S. F. (1898). The genesis of bitumens, as related to chemical geology. *Proceedings of the American Philosophical Society, 37*, 108–139.

Rémond de Saint-Mard, Toussaint. (1734). *Réfléxions sur la poësie en général, sur l'églogue, sur la fable, sur l'elégie, sur la satire, sur l'ode et sur les autres petits poëmes*. La Haye: C. de Rogissart & Soeurs.

Rowe, R. (1999). The aesthetics of interactive music systems. *Contemporary Music Review, 18*, 83–87.

Schroeder, M. (2009). *Number theory in science and communication: With applications in cryptography, physics, digital information, computing, and self-similarity* (5th ed.). Berlin: Springer.

Sebestik, J. (1983). The rise of the technological science. *History and Technology, 1*, 25–43.

Sirridge, M. (1988). Robert Kilwardby as 'scientific grammarian'. *Histoire Épistémologie Langage, 10*, 7–28.

Steele, R. B. (1900). The Greek in Cicero's Epistles. *American Journal of Philology, 21*, 387–410.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Tatarkiewicz, W. (1963). Classification of arts in antiquity. *Journal of the History of Ideas, 24*, 231–240.

Tonelli, G. (1975). The problem of the classification of the sciences in Kant's time. *Rivista Critica di Storia Della Filosofia, 30*, 243–294.

Tulley, R. J. (2008). Is there techne in my logos? On the origins and evolution of the ideographic term-technology. *International Journal of Technology, Knowledge and Society, 4*, 93–104.

Van Den Hoven, B. (1996). *Work in ancient and medieval thought: Ancient philosophers, medieval monks and theologians and their concept of work, occupations and technology* (Dutch Monographs on Ancient History and Archaeology, Vol. 14). Amsterdam: Gieben.

Vasari, G. ([1550] 1855). *Lives of the most eminent painters, sculptors, and architects*, vol. 1. (Translation by Mrs. J. Foster). London: Henry G. Bohn.

Weisheipl, J. A. (1965). Classification of the sciences in medieval thought. *Mediaeval Studies, 27*, 54–90.

Whitney, E. (1990). Paradise restored: The mechanical arts from antiquity through the thirteenth century. *Transactions of the American Philosophical Society, 80*, 1–169.

Young, J. O. (2015). The ancient and modern system of the arts. *British Journal of Aesthetics, 55*, 1–17.

Sven Ove Hansson is professor in philosophy at the Department of Philosophy and History, Royal Institute of Technology, Stockholm. He is editor-in-chief of *Theoria* and of the two book series *Philosophy, Technology and Society* (Rowman & Littlefield International) and *Outstanding Contributions to Logic* (Springer). His research areas include moral theory, value theory, philosophy of technology, philosophy of science, epistemology, and logic. He is the author of well over 300 articles in international refereed journals. His recent books include *The Ethics of Risk* (Palgrave Macmillan 2013) and *The Role of Technology in Science. Philosophical Perspectives* (Springer 2015). He is Past President of the Society for Philosophy and Technology.

Chapter 5
The Future of Philosophy: A Manifesto

Joseph C. Pitt

Abstract The future of philosophy is the philosophy of technology. It is argued that, using Wilfrid Sellars' aim for philosophy as "seeing how things in the broadest possible sense hang together in the broadest possible sense", contemporary philosophy is nothing more than a fragmented set of abstract and irrelevant activities. Philosophy, it is suggested should be about mankind interacting with the world, which is, on my account, the nature of technology. The role of philosophy should be to help us accomplish those interactions in a thoughtful and productive manner. The philosopher should be seen as part of a team of individuals seeking to accomplish something – she is a critical facilitator – Socrates reborn. To make philosophy a useful feature of the contemporary intellectual scene, we must disengage from minor analytic exercises that have little or no bearing on one another or the world and try to understand mankind interacting with the world, which would be to do philosophy of technology. To accomplish this goal I suggest that we reject the traditional taxonomy we appeal to when trying to make sense of philosophy. Instead of the old division of the field into Epistemology, Metaphysics, Value Theory, History of Philosophy, and Logic and Philosophy of Science, I urge that we first see the aim of philosophy to be assisting humankind to make their way in the world. The point is successful action. The traditional areas of philosophy and numerous subfields that don't fit easily into the traditional taxonomy such as aesthetics, philosophy of law and philosophy of technology, are now seen as areas to appeal for assistance in achieving the proposed goal. Every area of philosophical interest should be appealed to as a tool, not as a specialized area of "research".

Keywords Philosophy of technology • Philosophy as critical assistance • Philosophical taxonomy • Instruments • Eternal questions

---

J.C. Pitt (✉)
Virginia Tech, Blacksburg, VA, USA
e-mail: jcpitt@vt.edu

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_5

5.1 Introduction: The Nature of Philosophical Problems

Unbeknownst to me, over the years I have launched a two-prong attack on the practice and understanding of the analytic tradition in philosophy. This only first became clear when I was asked to contribute to the original empirical turn in the philosophy of technology project. Returning to the project in its present form gives me the opportunity to put together a number of ideas about Philosophy and the philosophy of technology that have been bouncing around in the back of my head for some time.

First, I will approach the very idea of what constitutes a philosophical problem. Second, I will investigate what I will call the taxonomy of philosophy. Fundamentally the idea I am pursuing is that it makes no sense to think of philosophy as an activity primarily engaged with eternal questions and issues. Third, I will argue that the proper subject matter of philosophy is man acting in the world. This is a topic that has been addressed by numerous European and South American authors, but somehow largely has been expunged from the Anglo-American analytic tradition and it is to that tradition that I am directing my remarks.

Let's start with the nature of philosophical problems. We have all been told from our first philosophy course or maybe even earlier that philosophical problems should be understood as eternal questions every generation asks and needs to answer. Furthermore, these questions mean the same thing over the ages. These are questions like "What is real?", "What is the good life?", What is knowledge?", "What is good reasoning?", and, as it happens, they pick out the basic areas of philosophy: metaphysics, value theory, epistemology, and logic.

I am going to suggest that this picture is incorrect, i.e., that there are no eternal questions for philosophers to puzzle over. There *may* be a tradition of asking questions that fall within a particular domain of inquiry, however, the questions do not remain eternally embalmed in Plato's heaven.

How did I come to this conclusion? As usual, I arrived at this point through my teaching. When teaching ancient western philosophy I try to get my student to understand the world of fourth century BC Athens. When teaching David Hume, I try to get my students to understand the world of eighteenth century AD England, Scotland and France. Reflecting on how I teach, I found myself asking the following question: What would a satisfactory answer to these so-called eternal questions look like in different times? Thus, what would an answer to the question "What is the good life?" look like in Socrates' Athens, where women did not have the vote and slavery was a standard practice as opposed to today, where we are struggling to understand the depths of institutional and social blockades to treating women and minorities as fully equal members of society and where slavery is simply abhorrent? An acceptable answer to that question will vary by place and time. Thus, I conclude, the question cannot mean the same thing over time.

Consider the alleged central question of epistemology, "What is knowledge?" This too will have a different acceptable answer in an age where the source of knowledge is unaided sense experience as opposed to our contemporary era where complicated instruments, electronics, and computers play such a key role. And it might be, no, should be argued, that analytic metaphysics is irrelevant today since the question of what is real is a question to be answered by physics, assuming one is some kind of scientific realist. Logic too is slipping away from philosophy as the inner workings of computers and the search for artificial intelligence require new and more powerful forms of reasoning than we philosophers deal with.

Let's assume the point is made – philosophical questions are not eternal – the same words may keep appearing, but what constitutes an adequate answer to the question keeps changing – hence the meaning of the question must change.

What about the areas of philosophical inquiry, epistemology, ethics, etc.? I have a problem with that as well. My difficulties here began with an attempt to figure out what philosophy really is. Is it a subject matter, a method, the history of what has been called philosophy? Let's look at that last one – its history. As the year 2000 approached everyone was making lists – 100 greatest novels of the twentieth century, etc. So since I was doing a lot of international conferencing at that time, I thought I would compile a list of the ten greatest philosophers of the twentieth century by asking my colleagues who they thought ought to be on the list – I asked Asian colleagues, European colleagues, South American, South African, Australian, etc. The final list was not what I hoped for, even though I didn't know what I was hoping for at the time. That final list did not include Mahatma Gandhi, nor Martin Luther King, nor Simone de Beauvoir. At the top of the list were Martin Heidegger, John Dewey, and W.V.O. Quine!!! So it appears that for analytic philosophers, people whose words and ideas actually made a difference and changed the world aren't philosophers.

The problem here, and believe me there is a problem, is that today we really don't know what philosophy is supposed to be nor what it is supposed to accomplish. Now if this were just me ranting, you should probably head down town. But many others are also dissatisfied with what philosophy has become: an ever increasing irrelevant search for minute answers to equally irrelevant questions. And there have been efforts to change that. Let me take as an example what has happened in the Philosophy of Technology. In an effort to escape charges of irrelevance, a group of us, lead by the Dutch philosophers from the Technical University of Delft launched a program in the late 1990s they called The Empirical Turn in the Philosophy of Technology. But, I will not just look at the empirical turn, I will use that venture as a platform to launch a total revision of our understanding of what philosophy is for and hence what its job is.

In many ways the empirical turn in the philosophy of technology (see Kroes and Meijers 2000) was very much like the turn to the special sciences in the philosophy of science in the second half of the twentieth century. As we paid more and more attention to the philosophy of physics, biology and now chemistry among others, the need to know the science was manifest and the abstract discussions in "general" philosophy of science became less and less relevant. For example, a general theory of scientific explanation becomes useless when we discover that explanations in physics differ from those in biology. The philosophical problems of the special sciences became far more interesting and relevant than abstract discussions because, as it turned out, the issues in the special sciences were increasing used as counterexamples to the claims of the general philosophy of science. Thus, the notion of evidence varies from science to science due to the fact that the evidence gathering instruments differ. Likewise, the empirical turn pushed us to know the technologies we were talking about, leaving abstract discussions of why Technology (with the capitol T) is evil, for instance, less and less interesting and irrelevant. Technologies have real world effects and knowing how that all works is crucial to knowing what to do with our technologies.

By taking the empirical turn we were forced to look at the things we can do with our technologies. We also began explorations of how individual technologies augment human capacities and raise new ethical and social issues.

5.2 Philosophy of Technology: Humans Acting in the World

What I will now propose and defend is the idea that all the areas of philosophy should be subsumed under the banner of the philosophy of technology, where the philosophy of technology is understood as the efforts to act successfully in the world. I am moved to this view by something Wilfrid Sellars said:

> The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term. Under 'things in the broadest possible sense' I include such radically different items as not only 'cabbages and kings', but numbers and duties, possibilities and finger snaps, aesthetic experience and death. To achieve success in philosophy would be, to use a contemporary turn of phrase, to 'know one's way around' with respect to all these things, not in that unreflective way in which the centipede of the story knew its way around before it faced the question, 'how do I walk?', but in that reflective way which means that no intellectual holds are barred. (Sellars 1963, p. 1)

and I will add to that "in order for us to move around in and manipulate the world successfully." This little addition however, has major ramification for Sellars' account, for it undercuts his aim of seeing how things in the broadest possible sense hang together. Contrary to Sellars, Philosophy is not this look at the world as a whole from a mountain top. Philosophy is the attempt to understand how what we do now will affect what we do tomorrow and tomorrow and tomorrow. Or to paraphrase Karl Marx, the point is not to talk about the world but to change it. Philosophy, in short, is nothing if not an aid to action. Philosophy of technology is (and if not, should be) fundamentally about human beings acting in the world (See Pitt 2000). It is in that context that we encounter epistemological, metaphysical, and ethical issues. The need to act comes first and then the more constrained philosophical concerns have the opportunity to raise their head, assuming we do as Sellars enjoins us, to think reflectively, just not too reflectively. The charges I have leveled against traditional analytic philosophy might lead some to the conclusion that I believe there are no philosophers in that mode concerned with a genuine connection between, for lack of a better way to phrase it, thought and action. This is not true.

The difficulty has been to find a mechanism to concentrate on. The one philosophical stance that actively seeks to make this connection is pragmatism. Beginning with C.S. Peirce (and before him, David Hume (see Pitt 2005)), whose main theme was the method to eliminate doubt, through Nicholas Rescher, the pragmatists have concentrated on acting in the world.

Therefore, in this chapter I propose to stop approaching philosophy as a set of independent areas that can be ordered in a taxonomy. Under the general rubric of the philosophy of technology reconceived as a philosophical analysis of humans acting in the world, I propose that we concentrate on diagnosing the various philosophical concerns that arise in the context of undertaking our real world projects as part of a team. In this view we cease to think of ourselves as metaphysicians or epistemologist, but as philosophers who can help identify metaphysical or epistemological issues as they arise while we do our work together with engineers, geologists, social planners, etc. The idea here is not to see the traditional areas of philosophy as separate areas of research in their own right, but as problem areas that need to be identified and dealt with in specific contexts.

What Sellars proposes as the aim of philosophy, to see how it all hangs together, cannot be achieved if we don't first have some sense of how the various components of philosophy itself hang together, for the ways we relate the various components of the world and human endeavors will be a function of the philosophical assumptions we bring to the party. Thus, how you see the fruits of research in physics will be a direct function of your views regarding the metaphysics of science. Further, you can't provide a systematic explanation of how it all hangs together if your philosophical view is itself fragmented. If you don't understand how your metaphysical views affect your epistemological endorsements, then there is no way to see how it all hangs together. The wrong way to do this is to specialize in one field, say, metaphysics. How can you lay claim to being a philosopher if all you do is worry about one small set of abstract and irrelevant problems in one small area of human thought? The philosopher's job is help us work through the real world implications for a proposed plan of action, not whether or not to be is to be the value of a bound variable (with apologies to Quine).

Let me be clear, I am not proposing a return to the grand metaphysical schemes of the nineteenth century. What I am proposing is a reorientation. When we introduce students to the world of philosophy we often tell them it consists of roughly five areas: epistemology, metaphysics, value theory, the history of philosophy, and logic and philosophy of science. But when pressed we have a hard time making sense of this overly simple taxonomy. For example, why is the philosophy of science a separate domain – doesn't it belong under Epistemology, and why is it lumped together with logic? Yes, it does seem like the odd man out, but there are sticky metaphysical questions about the reality of scientific objects and questions about method that, it has been argued, warrant separate treatment. What about aesthetics, the philosophy of law, and the philosophy of technology? How do they fit in? So as to appear more systematic than we really are, we have brushed aesthetics under the value theory rubric. The philosophy of law has also been relegated to value theory because laws, it is said, embody values. But the law is also a force for social change, and is itself a complex technology. So philosophy of law can be taken as part of philosophy of technology. When we undertake a legal action we engage in activities which affect the lives and fortunes of many others. The philosopher's role is to assist in ferreting out the implications of this or that legal move or attempt to change the legal system. We should be working with lawyers, judges, plaintiffs, and legislators to help determine the best path of action. And the philosophy of technology itself….? Well, in the years immediately following Heidegger, Philosophy of Technology would probably also have been considered part of value theory, but then given the growing closeness of science and technology in the minds of many, it was assumed to be part of the philosophy of science, but it is not, since the sciences and our technologies have little in common except that the sciences use technologies and our technologies sometime rely on scientific principles or discoveries and I am increasingly unsure that the philosophy of science requires a separate place in the taxonomy. If anything, philosophy of science should be subsumed under the philosophy of technology. We can't do science without the tools of mathematics and the means to conduct experiments. In short, science is technology-dependent and, thus, our views regarding science will be heavily influenced by our philosophy of technology. Most of us who work in the field think the philosophy of technology is a legitimate independent area of philosophy. But how can we justify this assumption?

When we turn to aesthetics it is not clear why it became lumped under value theory; for the simple question "What is art?" is not about value primarily, unless you arbitrarily decide to approach it that way, but it is certainly not necessary to do so – it can be a problem in metaphysics, for example when you ask what constitutes a work of art, is it the playing of the symphony, which when finished is gone, or is it the written score?

The problems here run deep. And to place blame, it all began with Aristotle's assumption that man is a rational animal (Nicomachean Ethics 1.13) instead of casting us as a social animal. More to the point, we are *homo faber*, man the maker. It was in the context of the social group that the first tools were produced. Simply put, once we move beyond the mistaken identification of technology with tools, we can appreciate the enormity, for example, of the development of agricultural practices as technologies for transforming the land. For with the development of agriculture it was possible to change from nomadic peoples following the herds with the seasons to settling down and building villages. This resulted in a number of transformations. The division of labor, beyond hunters/gathers, became possible and with it specialization, music, writing, and the further development of innovations to expand human expression and go beyond the efforts to simply secure survival. Aesthetics can thus be subsumed under philosophy of technology as well as we seek to understand what a better life means and how beauty contributes to the flourishing of the human spirit. Thus, how we live our lives becomes an aesthetic issue.

Of course this is written with a broad brush and the devil is in the details, for not all human development followed this pattern. For some there was no development beyond giving up the nomadic life, consider the aboriginal peoples of South America, Australia, and New Zealand. But in the cultures where we let the technologies lead us, we could also talk about human progress and wonder what the future will bring, something not possible in primitive societies content with the routines of day to day living.

What cannot be denied is that if we take the rational part of us as primary, we would have been someone's dinner long before we figured out how to survive by reason alone. Consider by way of example the opening scenes of Stanley Kubrick's film *2001*, *A Space Odyssey*. Kubrick is nothing less than brilliant in the insights he brings to the evolution of the human condition. We find two groups of apes/primitive men fighting over control of a water hole. At first there is a lot of screaming and grunting and pushing and it is apparent they don't really want to get truly physical. But then what appears to be the leader of one group picks up the femur of an animal that had died at the watering hole and using it instinctively as a club kills a member of the other tribe, thereby securing the water hole. In the scene he kills the other ape and looks at the femur and there appears to be a flash of understanding as to how this sort of thing can be used in the future as a tool (?) or weapon (?) and in exultation he tosses it into the air. It begins rotating slowly and the scene shifts to a space shuttle carrying passengers to a space station rotating around the moon, all to the sounds of the Blue Danube waltz.[1]

Kubrick gives us the connection between tools, warfare, technological development and the transformation of ape to man in an incredibly insightful presentation. It was important that he had two groups of apes in competition – the social group is primary, competition for survival is basic and the tool makes it possible.

So if we take that as our starting point, then we could argue that the tools are essential to the survival of the group. This places the philosophy of technology at the starting point in our efforts to form a coherent explanatory philosophical base from which to achieve, now somewhat limited, Sellars' aim for philosophy. For within the philosophy of technology, understood as understanding the relations between mankind and the world, we find all the questions of philosophy, perhaps slightly transformed. Let's see how this plays out.

5.3 From a Perennial to a Heraclitian Philosophy

Within the category of value theory we usually find ethics, metaethics, political philosophy, social philosophy and (wrongly) aesthetics. We also have a central question such as "What is the good life?". But to try to answer that question without understanding that it cannot be answered in the abstract by merely defining "good"

---

[1] It has been brought to my attention that some believe that the flash of insight the master ape had was the result of the aliens who planted the monolith; that the aliens in effect put that idea in his head. Needless to say, there is a lot of disagreement over how to interpret that film. Whether or not aliens helped, the key point is that the film's portrayal of early human behavior and the almost immediate clash with the future serves as a fruitful presentation of a powerful idea.

and "life" is to fail to see the bigger picture.[2] From our new perspective we see that it is not "What is the good life?" that needs answering – it is rather something like "What is required to *live* the good life?" Now we are talking about groups of humans interacting and being creative, seeking to minimize excess effort in favor of leisure and improvement. We can now ask in a meaningful way, "What do we need to live the good life?" This question takes us beyond traditional issues of ethics and the hypothetical best political system to actual material needs and how technologies affect the quality of life. We begin to see how our technologies are integral to our way of life and how they can contribute both positively and negatively. Let us assume that it has been decided that the generation of electrical power is essential to improving our way of living and the best way to do this in current circumstances to build a dam that will allow us to generate electricity. Having decided this, we are immediately led to epistemological issues such as "Do we know what the consequences of building this dam will be?" And that requires that we know what is involved in building a dam and how it affects the local ecology and the ecology downstream (philosophy of science). This can lead us to the question of the very nature of a dam – what is it (metaphysics)? How does it differ from the water it seeks to contain? Are there fundamentally different things in the world? If so, how can we use them to our advantage? Who should we trust to give us the answers to these questions, i.e., who has the relevant knowledge, and how do they interact with our leaders and politicians (our leaders being the CEOs of major multi-national corporations and our politicians are their dups). This inevitably raises questions of the social impact of the dam and how people and their way of living will be affected. But this requires that we have a grasp of the kinds of thinking and reasoning that would be appropriate to dealing with these issues, now enter logic. The philosophical questions here fall out of the development of a technology as we seek to make sense of what we are doing – and it is the doing that sets it all in motion. The role of the philosopher seen from this perspective is to help the team of actors involved in this. And the kinds of philosophical issues that arise arise because of something we want to do or the kind of project we are engaged in. So, in an important sense, there is no fixed taxonomy for philosophy. It is rather that the way and order in which philosophical questions arise have to do with what we are trying to accomplish. It also follows that as philosophers, in the spirit of the empirical turn, we need to know a lot about a lot of things, especially how things work.

This approach makes a lot more sense than simply asking "What is the Good Life?", "What is Real?", and "What can I Know?"; assuming these questions can be answered in the abstract and then be of use. The fact of the matter is that the only philosophical questions that are of use are the ones we have refined to the point where they can be turned into empirical sciences such as physics, astronomy, linguistics, economics, political science, etc., and, hence, cease to be philosophical issues. The history of philosophy has been the history of spin-offs. The questions we

---

[2] The real point here is move ethics away from an actor-centered perspective to a group-centered perspective. The actor-centered views such as utilitarianism and deontology rarely if ever have anything to do with how we act. If anything, they are employed in a casuistic manner, after the fact.

are left with as philosophers remain incoherent, framed as they are in isolation, and the answers we provide are useless since they are so abstract they fail to make contact with social reality.

To take the empirical turn seriously and then take the next step we begin with the idea that we are social creatures living in a physical world and it is our job as philosophers to understand that relationship in a coherent and explanatory manner.

There is one final point to consider here. In the perennial philosophy as conceived by Leibniz and used as a whipping boy by Sellars, philosophical questions are eternal questions, asked by all reflective people through the ages. Further, the answers to these questions are also assumed to be eternal. But with a little reflection we realize that those assumptions can't be correct. For the answers to the eternal questions, if there be any, change over time, as noted above. But here is where the discussion seriously diverges from what was said earlier. These changes are propelled by technological innovations. To rely on one of my favorite examples (manuscript in progress, *Seeing Near and Far: A Heraclitian Philosophy of Science*), consider how the answer to the questions "Can I trust what I see?" changes over time. To begin with what it means to see something changes as we introduce technological innovation into the game. When Galileo turned his telescope towards the moon he "saw" things seeable by the naked eye. Yet the telescope enhanced that seeing, thereby expanding the notion of what it is to see something. Likewise, for the concept of observation. The concept of observation changes over time due to the introduction of novel technologies that make what it is to be a scientific observation something very different from naked-eye seeing. When the NASA Galileo probe sent pictures back from its visit to Jupiter it involved a very complicated process. First of all the "camera" was not the family camera. Second, multiple instruments, mostly computer controlled, hence lots of interactive computer programs, were needed to keep the probe oriented towards the earth and then send the "picture" homeward. A major assumption at play is that nothing happened to the picture as it traveled through space. Then the information that had been transmitted was collected by the array at Arecibo and using multiple computers and programs finally portrayed on a screen – voila', a scientific observation of Io! Making this observation was not exactly the same as looking through a hand held telescope at the moon. And yet we still accept it as an observation of Io – our concept of what constitutes a scientific observation has changed, forced by the technologies we use.

As our technologies become more sophisticated this sort of conceptual change will happen at a faster rate, forcing us to continuously rethink our conceptual structures and our own relations to world. So the answer to "what do I know" will change as will all the others answers to the so-called perennial questions. The truly difficult philosophical task will be, in the light of constant change, to continue to see how it all hangs together. In short, the empirical turn takes us down the Heraclitian road.

To take the empirical turn seriously and then take the next step we begin with the idea that we are social creatures living in a physical world and it is our job to understand that relationship in a coherent and explanatory manner. Given that starting point we have a different job from philosophers laboring in the shadows of the perennial philosophy and one that may actually be doable. But first we have to understand that whatever coherent account we come up with will be constantly changing as our technologies change, forcing new social arrangements, making new discoveries possible, and posing new questions of justice and virtue. Scientific change is fueled by technological innovation, as is social change. And so it seems that since our technologies and what we do with them define us, then we should see all philosophical discussion as part of the philosophy of technology.

References

Kroes, P., & Meijers, A. (Eds.). (2000). *The empirical turn in the philosophy of technology*. Amsterdam: JAI/Elsevier Science.

Pitt, J. C. (2000). *Thinking about technology: Philosophical foundations of the philosophy of technology*. New York: Seven Bridges Press. http://www.phil.vt.edu/Pitt/jpitt.html

Pitt, J. C. (2005). Hume and Peirce on belief, or, why belief should not be an epistemic category. *Transactions of the Charles S. Peirce Society, XLI*(2), 343–354.

Sellars, W. (1963). Philosophy and the scientific image of man. In *Science, perception and reality* (Chapter 1, pp. 1–28). London: Routledge and Kegan Paul.

Joseph C. Pitt AB College of William and Mary. M.A. and Ph.D. in Philosophy from University of Western Ontario. Professor of Philosophy, Virginia Tech. Founding Editor of *Perspectives on Science: Historical, Philosophical, Social*. Former editor of *Techné: Research in Philosophy and Technology.* Latest book: *Doing Philosophy of Technology.* Founding Director of the Center for the Study of Science in Science, which has transformed over the years into the Graduate Program in Science and Technology Studies. Major interests: the impact of technological innovation on scientific change. Current Projects: Editor of *The Routledge Companion to the Philosophy of Technology* (in progress); *Seeing Near and Far, a Heraclitean Philosophy of Science.*

Chapter 6
Science vs. Technology: Difference or Identity?

Ilkka Niiniluoto

Abstract It is argued in this paper that there is an important conceptual distinction between science and technology. As parts of human culture and society, science and technology exist today in a state of dynamic mutual interaction, but differences can be found in their aims, results, and patterns of development. Therefore, there are also significant differences in science policy and technology policy. This conclusion is at variance with the fashion of using the term "technoscience" in the STS-studies. Some critical comments are also given on the social constructivist treatment of "sociotechnology".

Keywords Applied research • Scientific realism • Sociotechnology • Technoscience

6.1 Creating and Blurring Distinctions

Philosophers, especially those belonging to the analytic tradition, are usually fond of making conceptual distinctions. As an activity, philosophy aims at clarity through the method of conceptual analysis. Examples of such important distinctions include matter-mind, object-subject, reality-appearance, truth-falsity, theory-practice, nature-culture, sex-gender.

Another trend is the attempt to question and to abolish such conceptual differences. American pragmatism (from John Dewey to Richard Rorty) and French postmodernism and deconstructionism (Jacques Derrida) are philosophical programs for blurring and abandoning binary oppositions.

Conceptual distinctions are not philosophically innocent, but typically involve or presuppose wholesale theoretical and even ideological frameworks. Defending and challenging, or creating and blurring, distinctions are two important aspects of philosophical investigation. But, on the other hand, the results of such investigations

---

I. Niiniluoto (✉)
University of Helsinki, Helsinki, Finland
e-mail: ilkka.niiniluoto@helsinki.fi

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_6

cannot be known in advance on the basis of some general program, but each issue has to be studied separately in a careful manner.

In particular, what is said above applies to the scientific study of science, technology, and society (STS). In 1962 the influential Frascati Handbook of the OECD consolidated two distinctions which have been widely used in science policy. First, it distinguished *research* ("the pursuit of new knowledge") and *development* (the use of results of research "to develop new products, methods, and means of production"). The roots of this R&D divide go back to Aristotle's division between *episteme* and *techne*. While *episteme* (Lat. *scientia*) means knowledge, or justified true beliefs expressible by propositions, *techne* is defined as a rational and stable habit of making or producing material objects (see *Nicomachean Ethics* VI, 4; 1140a1). This difference between scientific knowledge and productive arts is the basis of our standard distinction between science and technology. Secondly, the OECD Handbook made a distinction between two kinds of research: *basic research* (also called fundamental, curiosity-driven or blue skies research) seeks knowledge for its own sake "without the aim of specific application", while *applied research* (also called mission-oriented research) pursues "knowledge with the aim of obtaining a specific goal".

It is no wonder that the OECD terminology has been challenged in many ways. For example, the distinction between basic and applied research has been rejected by many scholars as obsolete (see e.g., Douglas 2014), while some others have still defended the importance of this division in some refined form (see Niiniluoto 1984, 1993; Sintonen 1990).[1]

Another example is the distinction between *science and technology*. The traditional conjunctive way of speaking suggests that science and technology are two different parts or sections of human activities. But it has become fashionable in the STS-studies to combine the two into the single term *technoscience*. Bruno Latour (1987, p. 29) tells that "in order to avoid endless 'science and technology' I forged this word".[2] The Society of the Social Study of Science (4S) has adopted this new term as the title of its Newsletter, to indicate that its scope includes what used to be called the sociology of science and the sociology of technology. But it is clear that "technoscience" is not only a shorthand notation for a longer phrase, but it aims at blurring an old distinction and thus constitutes a central and essential element of a new ideology about the subject matter and methods of science studies.

A similar strategy is followed by Wiebe Bijker and John Law (1992), who use the constructivist approach to deconstruct the science-society distinction. On the basis of their idea of a "seamless web", they introduce the term *sociotechnology*.

---

[1] Douglas gives an interesting account of the emergence of the ideological contrasts between pure and applied science in the nineteenth century, but she seems to forget that the "rhetoric invention of pure science" took place already in the ancient Greece. She concludes that scientific progress should be defined in terms of "the increased capacity to predict, control, manipulate, and intervene in various contexts". As Douglas does not make a difference between applied research (as pursuit of special kind of knowledge) and the applications of science (in control and problem-solving), she is not in fact relinquishing the pure-applied divide but rather the science-technology division.

[2] In fact the first who used this term (in French) was the philosopher Gaston Bachelard in 1953.

Proceeding along these lines, one might suggest that the two neologisms are further combined into *sociotechnoscience*!

The current trend in the field seems to advocate the principle: what pragmatists, deconstructionists, and sociologists of science have united, an analytic philosopher shall not divorce or separate! But, even though it may appear old-fashioned, I still believe in the value of some distinctions in the STS-studies.

More precisely, in this paper, I defend the view that there is an important conceptual difference between science and technology. This thesis does not imply that the distinction is absolutely clear cut: day differs from night, even though there are unsharp borderline cases (twilight). And it certainly does not mean that science and technology have nothing to do with each other, or that their relations are historically constant. A significant distinction may be "formal" in the scholastic sense: science and technology usually occur together in social reality and interact with each other in modern societies, just as length and weight are two distinct but coexisting aspects of physical objects.

6.2 Five Models for the Science-Technology Relationship

Don Ihde (1979) has made an illuminating comparison between the science-technology distinction and the alternative solutions to the classical mind-body problem.[3] In metaphysics, there are five main models of the relationship between mind (spirit) and body (matter):

1. Idealism: mind is ontologically primary to body.
2. Materialism: body is ontologically primary to mind.
3. Identity: mind and body are the same.
4. Parallelism: mind and body are causally independent but parallel processes.
5. Interactionism: mind and body are ontologically independent but causally interacting.

The first three doctrines are monistic, as they assume only one basic substance. Excluding here the radical eliminativist versions, ontological primacy means that (1) the existence of bodies depends on the existence of minds, or that (2) minds cannot exist without bodies. The formulation of idealism can be reductive (bodies are reducible to minds) or emergentist (bodies are results or products of minds), and similarly for materialism.

The last two doctrines are dualistic, as they assume two ontologically independent substances. Five says that mind and body causally influence each other, while four denies this but still claims them to behave in some correlated fashion.

---

[3] I have followed Ihde's presentation in my first papers on the philosophy of technology (see Niiniluoto 1984, Ch. 12). Kusch (1996) has used the same idea in his discussion of the cognitive-social distinction. For the mind-body problem, where my sympathies are for emergent materialism, see Niiniluoto (1994).

Following this schema of alternatives, but replacing “mind” with “science” and “body” with “technology”, we obtain five different positions:

1. Technology is reducible to science, or technology depends ontologically on science.
2. Science is reducible to technology, or science depends ontologically on technology.
3. Science and technology are identical.
4. Science and technology are ontologically and causally independent.
5. Science and technology are ontologically independent but in causal interaction.

Thesis (1) is implied by the standard view which defines technology as applied science or as the application of science. This view, which can be found in many dictionaries of English language, gains some support from etymology: “technology” is the *logos* (doctrine, learning) of *techne* (art, skill, technique) (see Mitcham 1994). This seems to suggest that technology is a special branch of human knowledge (Lat. *scientia*) (see Bunge 1966).

But in English “technology” may also mean collections of tools and machinery, and the art of designing and using such material artefacts to produce other artefacts. Aristotle, who distinguished *praxis* (activity which includes its own purpose) and *poiesis* (making or production), defined *techne* as a rational and stable habit of making or producing. Ihde (1983) uses the word “technics” essentially in this sense.[4] Many languages prefer terms derived directly from *techne* (e.g., “Technik” in German, “téchnica” in Spanish, “tekniikka” in Finnish) to words including *logos*.

Plato and Aristotle recognized that rational skills presuppose or contain background knowledge in different degrees. But this knowledge-ladenness of skills does imply that art is nothing but knowledge. Following Ryle (1949), it has been argued that specific technologies – as professions, practices, and arts – involve *know how* which cannot always be reduced to propositional *know that*. The emergence of the philosophy of technology in the 1960s, as a field of analytic philosophy independent of the philosophy of science, was mainly based on the observation that technology should not be identified with applied science (see Rapp 1974; Bugliarello and Doner 1979).

As Ihde (1979, 1983) convincingly observes, the thesis (1) is in conflict with the fact that technology has historical priority over science. As “tool-making animals” (Benjamin Franklin, Karl Marx), our ancestors have designed and used tools and artefacts at least for 3 million years. As systematic pursuit of knowledge, which presupposes the use of symbolic languages, science has existed only for 3000 years. Therefore, technology on the whole cannot be ontologically dependent on the existence of science, which is a latecomer in human culture.

Thesis (1) has still restricted validity in the sense that there exist technical artefacts that have been made possible only by the progress of science (e.g., nuclear

---

[4] For the Greeks the productive arts included also poetry. The word “technology” in the broad sense can cover, besides instrumental action or work with tools, also expressive action (e.g., play with toys and musical instruments).

bombs and reactors were built by using information provided by physical theories about atoms and radioactivity; similarly mobile phones and led lights have been designed on the basis of engineering sciences). Such *science-based* technology, which fulfils Francis Bacon's vision of knowledge yielding power, is called *development* by the OECD. But, historically speaking, all areas of successful technologies have not in fact been based on scientific theories. This is the case with some old parts of folk medicine, the industrial revolution of the eighteenth century (steam engine, spinning machine), military technology up to the late nineteenth century, and many patented inventions even today.

Thesis (2) is implied by the *instrumentalist* view which takes theories to be sophisticated conceptual tools of human practice, and thus science to be a tool of technology. Science is seen as a moment in the human endeavor to master nature. This view thus strengthens the historical priority of technology to its ontological priority over science. Ihde associates this doctrine with the "praxis philosophies" (pragmatism, Marxism, phenomenology, Heidegger).

Instrumentalism represents a technological conception of science, which takes science to be always governed by the "technical interest" to control reality for human purposes (in the sense of Habermas). This characterization may fit applied "design science" (cf. Niiniluoto 1993), which seeks lawlike and manipulable connections between means and ends. However, it is not adequate to basic research whose goal is true descriptive and explanatory information about reality, independently of practical applications.

Instrumentalism also fails to explain the historical fact that science was born in the ancient Greece as a theoretical activity of the philosophers of nature who wished to uncover the basic elements of reality by using their reason, without relying on old myths and religions. The connection between such theoretical science and practical action was largely unknown to the Greeks.[5] For Aristotle, the practical sciences included ethics and politics, and they were distinguished both from the theoretical sciences (such as mathematics, physics, and theology) and from the productive arts.

The *identity* thesis (3) treats science-and-technology as a single totality without distinction. Given the great temporal differences in their development, the idea of the original identity of science and technology is entirely implausible. But a more interesting version claims that science and technology have become identical in the modern age. In the stone age, there was still pure technics without science, and in the ancient Greece the philosophers of nature were engaged in theoretical science without technology. But through the Baconian scientification of technology, the instrumentation of scientific research, the emergence of Big Science and applied research, industrial laboratories, and science-based development, it may appear that science and technology have been fused into a new conglomerate. The term "technoscience" might be used as a name for this new unity.[6]

---

[5] The work of Archimedes on mechanics and hydrostatics had practical applications, but it was not "applied research" in the modern sense.

[6] Here I am using this term in an ontological sense to express the alleged "real" identity of science and technology. It should be noted that in the constructivist science studies the term "technoscience" is mainly used in the methodological sense, i.e., a sociologist should proceed without any

If the term “technology” is used in a broad sense which covers technical or engineering sciences, then the intersection of the areas of science and technology is of course non-empty. But this does not imply the identity thesis, since there still are branches of basic sciences (such as fundamental physics, biology, and sociology) which cannot be included among the technical sciences.

The *parallelist* thesis (4) has not found many supporters. It was defended by Derek de Solla Price (1965), who compared science and technology to two dancers who make similar movements by following the same tune (but not interacting with each other).

The *interactionist* view (5) claims that there are mutual causal influences between science and technology. In my view, this is the “dualistic” position that best explains the independent historical origins of technology and science. It admits that especially since the late nineteenth century there is an important overlap area, which includes science-based technology and instrumentally embodied research, but – in contrast to the idea of “technoscience” – even in this joint area it is still possible to conceptually distinguish the elements or aspects that are descendants of science and technology, respectively. (Similarly, some children resemble their father, some their mother.) Hence, against the identity thesis, it is still possible to distinguish science and technology from each other – even in those cases, where both are parts of the same research institute or project, or both are parts of the work of the same research group or individual researcher.

For example, for Lacey (2010) nanotechnology is a paradigm of what he calls “technoscience”: the use of advanced technology and instruments to gain knowledge about new possibilities that we can do and make, with the horizons of practical and industrial innovation, economic growth and competition. In nanotechnology, theoretical knowledge about physics and chemistry is used to develop new nanomaterials that have economically profitable industrial applications. But the results of nanotechnology include both knowledge and artefacts: scientific articles in specialized journals like *Nano Letters* and *Advanced Materials* and new material products.[7]

The forms of the science-technology interactions are historically changing. Today they are more intensive and variegated than ever.

Technology provides new instruments for scientific research (thermometers, telescopes, microscopes, chemical and medical laboratories, high energy accelerators, computers, etc.). A prominent example is the use of the Large Hadron Collider (LHC) at the CERN Laboratory to test the Standard Theory of matter and to hunt the Higgs particle. Technological practices and inventions create new research problems, theories, areas, and disciplines (e.g., steam engine and thermodynamics, farming and agricultural sciences, telephone and information theory, computer and computer science). Technology may also provide concepts and models that are used

---

initial assumption about the difference between science and technology. However, the problems with the ontological identification of science and technology carry over to the methodological distinction as well.

[7] Nanotechnology is used as an example of technoscience by Nordmann (2016).

in scientific thinking as metaphors or theoretical concepts (the world as a clock, the heart as a pump, the human mind as a Turing machine, etc.). Finally, technological progress indirectly influences science by fostering economic growth.

Causal influences from science to technology include innovation chains from basic research to applied research and development, i.e., science-based design and production of new tools and devices. Today, such applications are often based on innovation cycles, where researchers, engineers, designers, and the potential customers interact with each other. Such "strategic" or "mode 2 research" (Gibbons et al. 1994) combines multidisciplinary basic and applied research with demand- and user-driven development of new products (Veugelers et al. 2009). Science may also help to explain why artifacts and methods work.[8] The education of engineers and technicians is also influenced by scientific knowledge and scientific methods.

6.3 Realism, Instrumentalism, and Constructivism

Different views about the distinction between science and technology depend on more basic philosophical positions concerning reality, language, and human action.

On the basis of *scientific realism*, one can characterize basic or fundamental research in science in the following way (cf. Niiniluoto 1984, 1999): research is an activity of the members of the scientific community; the method of science is based upon interaction between the scientists and the objects of their study; by using the methods of science researchers produce knowledge; this body of knowledge is formulated in language as sentences, propositions, laws, or theories; the aim of knowledge is to represent or describe some aspect of reality; such representations should give true or at least *truthlike information about reality*. For a realist, truth should be explicated as a relation of correspondence between linguistic representations and reality.

According to the realist view, the reality as the object of scientific inquiry may include nature, human mind, culture, and society. In Popper's terms, it includes Worlds 1, 2, and 3. Thus, science may study aspects of reality which are mind-independent (World 1) or ontologically dependent on human social activity (World 3) (see Niiniluoto 2006). In particular, language and linguistically formulated items of human knowledge are parts of World 3, and the same is true of science as a social institution. As an activity within the world, scientific inquiry may influence or "disturb" the reality under investigation (e.g., measurement in quantum theory, interview methods in social science). But this reality is still pre-existing in the sense that it is not produced or constituted by the research process. All this is compatible with the realist view of knowledge as more or less truthlike representation.

---

[8] Bunge (1966) calls "pseudotechnologies" such branches of technology that cannot be explained by science. Examples could include some medical treatments and pedagogical doctrines. This should not be understood to imply that all instrumentally rational technologies have to be genetically science-based, since this would be contrary to historical facts.

The basic "epistemic utilities", defining the aim of scientific inquiry, are *truth* and *information*. An important function of basic research is to explain observable phenomena and regularities by means of laws and theories involving theoretical entities and mechanisms. Science seeks not only *knowledge that* but also *knowledge why*, so that *explanatory power* is a central epistemic utility. Applied research also aims at knowledge, but, besides truth and information, it requires results that are *useful* or *socially relevant* for some human purpose. For some applied sciences (e.g., meteorology) this purpose may be *prediction*, which helps us to prepare for contingent events in the future. A special form of applied research, *design science*, is characterized by the goal of finding knowledge that expresses "technical norms", i.e., relations between means and ends (see Niiniluoto 1993, 2014). Such conditional rules of action give us *know how* by promoting some human professions and technological activities – such as the manipulation and control of some natural or artificial system so that some desired goal is achieved. The truth or correctness of a technical norm depends on the existence of an appropriate causal relation between an action and its goal. Examples of design sciences in this sense include agricultural sciences, engineering sciences, clinical medicine, nursing science, and social policy studies.

On this realist view, technology differs from science in the following way: the technologists (e.g., engineers, craftsmen, artisans, designers, architects) use the methods of design to create new artefacts or tools; such artefacts are material entities or prototypes of such entities; usually the products of technology are not formulated in language, and they do not have truth values; the tools have a specific purpose of use; the use of tools opens new *possibilities* for human action.

Instead of truth and information, the technological artefacts should be evaluated by the value of the new possibilities that they open. The basic utilities of technology are then *effectivity* relative to the intended purpose of tools (e.g., the destroying power of arms, the ability of ships to carry passengers) (cf. Skolimowski 1966; Sahal 1981), and their *economical* value (or cost-effect-*efficiency*) in terms of the required resources and expected gains (cf. Elster 1983). Further, all artefacts can be evaluated on the basis of their *esthetic*, *ergonomical*, *ecological*, *ethical*, and *social* aspects. This is the task of Technology Assessment (TA) (cf. Durbin and Rapp 1983).[9]

The realist's division between science and technology can be challenged in several ways. The *instrumentalists*, or pragmatists in the broad sense (cf. Rescher 1977), may accept that scientific theories as systems of statements (or as networks of models) are different from material artefacts, but claim that nevertheless theories too are in some sense human-made artefacts.[10] It is indeed correct to point out that theories and models are human creations, but still they are "epistemic artefacts" (Knuuttila 2005) with the aim of giving truthful information about reality.

The pragmatist may go further by claiming that scientific theories are tools: their ultimate aim is to enhance human relations to the natural and social environment, and their value is to be measured by the practical gains of their applications

---

[9] I discuss technology assessment in more detail in Niiniluoto (1997).

[10] For a general account of artefacts, see Franssen et al. (2014).

(Douglas 2014). One way of formulating this view is to claim that science is a problem-solving rather than truth-seeking activity.[11]

The realist replies to the instrumentalist that the practical value of scientific theories is derived from their goodness as representations of reality. Truthlikeness is a more basic goal of science than any practical utility, since it serves to explain the ability of a theory to yield useful applications. The predictions or conditional rules of action are reliable to the extent that they are derivable from true or truthlike theories. Thus, the problem-solving capacity of a theory presupposes some degree of success in truth-seeking (cf. Niiniluoto 1999).

Another objection to realism arises from the *constructivist* approach. The classical exposition of this view, Latour and Woolgar (1979/1986), argues that scientific facts and theoretical entities are social constructions. Reality is a "consequence" of scientific work, not its "cause", i.e., reality is the result of a process of negotiation and settlement of opinion within a local laboratory community.[12]

If the constructivists only mean that accepted scientific hypotheses and theories are social constructions, results of social negotiations and "closure" of controversy, their view is trivially compatible with both realism and instrumentalism. But if they further claim that scientific disputes are settled by appealing to the personal and social interests of the participants, and methodological considerations of truth and justification play no role in the process, their position is incompatible with epistemological realism.

However, the constructivists have also insisted that nature is a social construction. If a methodological rather than an ontological jargon is used, nature should be treated as a social construction. Bruno Latour's *Science in Action* (1987) claims that no sound divide can be made between scientific facts (e.g., the Watson-Crick model of DNA) and technological artefacts (e.g., Diesel motor, computer). In later work, this view has been generalized to the symmetry thesis that nature and society are both results of human scientific-technological activities (cf. Jasanoff et al. 1995).

Following Charles Peirce, the realist can argue against the constructivist that we obtain an incomplete and misleading picture of the research process, if the causal interaction with external reality is ignored (see Niiniluoto 1999, Ch. 9.3). Constructivism reverses the natural order of explanation: the existence of real things and facts "out there", together with the basic nature of belief formation by the scientific method, explains the consensus among scientists, not vice versa.

However, as artefacts are human-made, social constructivism may seem to be much more promising in the context of technology. But even in this field the constructivists have formulated views that appear problematic from the realist

---

[11] This is Larry Laudan's (1977) formulation. Unlike typical instrumentalists, Laudan admits that scientific theories have truth values in the realist sense, but he thinks that truth is a utopian goal in science and therefore irrelevant for scientific progress. For the distinction between cognitive problems and problems of action, see Niiniluoto (1984, Ch. 11).

[12] The description of social construction by means of negotiation and eventual consensus is different from the material construction of new artefacts (such as radioactive substances and synthetic materials) in laboratories.

perspective. Latour's (1987) "First principle" claims that the qualities of facts and machines are "a consequence, not a cause, of a collective action", and his "Rule 2" states that to determine "the efficiency or perfection of a mechanism" we should not "look for their *intrinsic* qualities but at all the transformations they undergo *later* in the hands of others". Earlier, Pinch and Bijker (1984) argued that the working or non-working of an artefact is not an explanation of its success or failure: such working is not an intrinsic property but rather socially constructed. Thus, machines work because they have been accepted by relevant social groups, not vice versa (see Bijker 1995, p. 270).

In my view, it is correct to stress that technological change is contingent and socially shaped (cf. Bijker and Law 1992): even though artefacts often create new social needs, as World 3 entities they are also formed to satisfy our interests. It is up to us to design and build machines so that they "work" well. When an artefact has been built, we are always free to change its properties later. The color of a car, or the efficiency of its engine, are in this sense not "intrinsic" or permanent properties. But, on the other hand, at any moment the artefact possesses such properties or functions in an objective way, and they also explain the "working" of the artefact (e.g., the maximum speed of the car, how it appeals to buyers).

6.4 A Difference in Dynamics?

To speak about the "seamless web" of "sociotechnology" – or even "sociotechnoscience" – has the virtue of warning that science and technology should not be studied in an "atomistic" way, but their involvement and interaction with social factors should be acknowledged in STS-studies. The use of the term "technoscience" is legitimate, if it is intended to remind us that today science and technology are treated together in most policy frameworks of R&D.[13]

The trend of combining science and technology can be illustrated by administrative developments in Finland.[14] It used to be the case that the main decisions about science policy, such as the funding of basic research in the universities, were made by the Ministry of Education. Technology policy, with its special interest in science-based development within technological universities and private research laboratories, was the domain of the Ministry of Commerce and Industry. Besides the funding of basic research by the Academy of Finland, a new funding agency for technology Tekes was established in 1983. In 1984 the former Science Policy Council was changed to a new Science and Technology Policy Council, and universities were encouraged to engage in partnerships with industry by establishing science parks to promote startup companies. In the totality of R&D funding in Finland, about 30 % has come from public sources and 70 % from private companies. Since the

---

[13] In his later work, Ihde has used the term "technoscience". See Ihde (2003).

[14] Similar stories can be told about other countries as well.

mid-1990s, the rhetoric for policy decisions about both science and technology have been based on the view of research and development as parts of the "national innovation system", which has the purpose of improving the economic competitiveness of Finland: *innovation* is here defined as "a novel good or service" or "an exploited competence-based competitive asset" (Veugelers et al. 2009). Tekes, now under the new Ministry of Employment and the Economy, changed its name to the Finnish Agency for Technology and Innovation, and the Science and Technology Policy Council is now Research and Innovation Council. In this way, science policy has step by step been subordinated under the instrumentalist or technological conception of science as a tool of economy.[15]

Even though the political decision-makers have – at least so far – been wise enough to continue to support basic research, the marriage of science and technology policy is potentially harmful to science. Investment in strategic research and innovations may bring about solutions to wicked problems and short-term profits, but neglect of independent fundamental research weakens the scientific community, the universities, and the economy in the long run. What is more, both the instrumentalist conception (with its reduction of science to technology) and the current STS-approach to "technoscience" (with its methodological identification of science and technology) seem to support such administrative solutions.

The dynamic models of scientific change (Kuhn, Popper, Feyerabend, Lakatos, Toulmin, Laudan, and others) became a hot issue in the philosophy of science in the 1960s and 1970s, and their relevance to science policy were also debated (cf. Niiniluoto 1984). Similar questions about technological change can be formulated in a fruitful way: internalism vs. externalism, qualitative vs. quantitative indicators, black box vs. content, revolution vs. cumulation, technocratic vs. democratic, inner logic or external control, determinism vs. voluntarism (cf. Ellul 1964; Winner 1977; Bugliarello and Doner 1979; Elster 1983; Laudan 1984; Sahal 1981; Niiniluoto 1990). Such models suggest interesting structural patterns that seem to be similar on the surface level of the development of scientific knowledge (e.g., Newtonian physics) and technological projects (e.g., cars, semiconductors) – for example, one may compare Kuhn's paradigm-based normal science, Lakatos's notion of a research programme, and Dosi's (1982) notion of technological trajectory.

However, the underlying dynamics seems to be quite different in the cases of science and technology, and this implies also a crucial distinction between the principles of science policy and technology policy (cf. Niiniluoto 1997). The decision to allocate funds for high-energy physics belongs to science policy, and here the best advice is obtained by scientific experts using peer review methods.[16] But the

---

[15] On the European scale, the EU framework programs had the aim of strengthening the economic competitiveness of Europe. The funding of ERC still leaves room for free basic research, since it is based solely on considerations of excellence and quality.

[16] In strategic research in the mode 2, there is room for the advice of potential users of scientific information, but this is much more restricted than the possibilities of using user-driven methods (such as consensus conferences) in technology assessment (see Shrader-Frechette 1985).

assessment of the credibility of the theory of quarks belongs to science itself. Such a theory should be accepted on the basis of its explanatory power and the experimental evidence supporting it, and such features depend on the nature of mind-independent reality. It is not up to us to decide whether there are gravitational forces or quarks in nature – and whether theories about such entities are true or false.

On the other hand, the decision to build the fifth nuclear power plant in Finland is decided in the Parliament. Small-scale decisions about the use of technological artefacts (such as clothes, furniture, household machines) are made by consumers in their everyday life. For some types of products and tools, there are social restrictions and controls (e.g., guns, medicine). The difference to scientific theories is clear and distinct: it is up to us to decide what artificial technological devices we wish to be created, produced, manufactured, and used in our society. For this purpose, we should develop democratic procedures of assessing and controlling technological change.

6.5 Conclusion

We have argued in Sect. 6.2 that science and technology are in interaction without being identical. This means that philosophy of science and philosophy of technology should likewise be in interaction without being reducible to each other. These disciplines have separate agendas which reflect the differences in the aims, results, patterns of development, and policy decisions of science and technology (Sects. 6.3 and 6.4). In particular, the key issue for philosophy of science is the production of truthlike knowledge about reality, while philosophy of technology should investigate the special ontological nature of artefacts – ranging from specific tools (like screwdrivers) to large-scale socio-technical systems (like cities). Such studies should acknowledge the differences in the value standards for assessing knowledge claims and new technologies.

On the other hand, within the mutual collaboration of these disciplines, philosophers of science are expected to develop accounts of new modes of applied research which are useful for the science-based design of artefacts and user-driven innovations. Such innovations range from mobile phones and intelligent robots to social services in the public sector. Philosophers of technology should follow the "empirical turn" of philosophy of science by showing how new technologies and engineering practices create instruments which help to explore and test scientific hypotheses. Besides various kinds of measurement devices and detectors, such instruments include computer methods which allow analogical inferences from idealized models to real target systems (Niiniluoto 2013).

References

Bijker, W. E. (1995). *Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change*. Cambridge, MA: MIT Press.

Bijker, W. E., & Law, J. (1992). *Shaping technology/building society: Studies in sociotechnical change*. Cambridge, MA: MIT Press.

Bugliarello, G., & Doner, D. B. (Eds.). (1979). *The history and philosophy of technology*. Urbana: University of Illinois Press.

Bunge, M. (1966). Technology as applied science. *Technology and Culture, 7*, 329–349. (Reprinted in Rapp 1974, pp. 19–36).

De Solla Price, D. J. (1965). Is technology historically independent of science? A study in statistical historiography. *Technology and Culture, 6*, 553–568.

Dosi, G. (1982). Technological paradigms and technological trajectories. *Research Policy, 11*, 147–162.

Douglas, H. (2014). Pure science and the problem of progress. *Studies in History and Philosophy of Science, 46*, 55–63.

Durbin, P., & Rapp, F. (Eds.). (1983). *Philosophy and technology*. Dordrecht: D. Reidel.

Ellul, J. (1964). *The technological society*. New York: Alfred A. Knopf.

Elster, J. (1983). *Explaining technical change*. Cambridge: Cambridge University Press.

Franssen, M., Kroes, P., Reydon, T., & Vermaas, P. (Eds.). (2014). *Artefact kinds: Ontology and the human-made world*. Dordrecht: Springer.

Gibbons, M., et al. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Sage.

Ihde, D. (1979). *Technics and praxis*. Dordrecht: D. Reidel.

Ihde, D. (1983). *Existential technics*. Albany: State University of New York Press.

Ihde, D. (2003). *Chasing technoscience: Matrix for materiality*. Bloomington: Indiana University Press.

Jasanoff, S., Markle, G. E., Petersen, J. C., & Pinch, T. (Eds.). (1995). *Handbook of science and technology studies*. London: Sage.

Knuuttila, T. (2005). Models, representation, and mediation. *Philosophy of Science, 72*, 1260–1271.

Kusch, M. (1996). Sociophilosophy and the sociology of philosophical knowledge. In S. Knuuttila & I. Niiniluoto (Eds.), *Methods of philosophy and the history of philosophy* (Acta Philosophica Fennica, Vol. 61, pp. 83–97). Helsinki: The Philosophical Society of Finland.

Lacey, H. (2010). Reflections on science and technoscience. *Scientiae Studia, 10*, 103–128.

Latour, B. (1987). *Science in action*. Milton Keynes: Open University Press.

Latour, B., & Woolgar, S. (1979/1986). *Laboratory life: The social construction of scientific facts*. Princeton: Princeton University Press.

Laudan, L. (1977). *Progress and its problems*. London: Routledge and Kegan Paul.

Laudan, R. (Ed.). (1984). *The nature of technological knowledge: Are models of scientific change relevant?* Dordrecht: D. Reidel.

Mitcham, C. (1994). *Thinking through technology: The path between engineering and philosophy*. Chicago: University of Chicago Press.

Niiniluoto, I. (1984). *Is science progressive?* Dordrecht: D. Reidel.

Niiniluoto, I. (1990). Should technological imperatives be obeyed? *International Studies in the Philosophy of Science, 4*, 181–187.

Niiniluoto, I. (1993). The aim and structure of applied research. *Erkenntnis, 38*, 1–21.

Niiniluoto, I. (1994). Scientific realism and the problem of consciousness. In A. Revonsuo & M. Kamppinen (Eds.), *Consciousness in philosophy and cognitive neuroscience* (pp. 33–54). Hillsdale: Lawrence Erbaum Associates.

Niiniluoto, I. (1997). Technology policy in a democratic state. In S. Hellsten, M. Kopperi, & O. Loukola (Eds.), *Taking the liberal challenge seriously* (pp. 192–204). Aldershot: Ashgate.

Niiniluoto, I. (1999). *Critical scientific realism*. Oxford: Oxford University Press.

Niiniluoto, I. (2006). World 3: A critical defence. In I. Jarvie, K. Milford, & D. Miller (Eds.), *Karl Popper: A centenary assessment II* (pp. 59–69). Aldershot: Ashgate.

Niiniluoto, I. (2013). Models, simulations, and analogical inference. In V. Karakostas & D. Dieks (Eds.), *EPSA 11 perspectives and foundational problems in philosophy of science* (pp. 19–28). Cham: Springer.

Niiniluoto, I. (2014). Values in design sciences. *Studies in History and Philosophy of Science, 46*, 11–15.

Nordmann, A. (2016). Changing perspectives: From the experimental to the technological turn in philosophies of science and technology. In M. Franssen, P. E. Vermaas, P. Kroes, & A. Meijers (Eds.), *Philosophy of technology after the empirical turn*. Cham: Springer.

Pinch, T. J., & Bijker, W. E. (1984). 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.

Rapp, F. (Ed.). (1974). *Contributions to a philosophy of technology*. Dordrecht: D. Reidel.

Rescher, N. (1977). *Methodological pragmatism*. Oxford: Blackwell.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Sahal, D. (1981). *Patterns of technological innovation*. Reading: Addison-Wesley.

Shrader-Frechette, K. (1985). Technology assessment, expert disagreement and democratic procedures. In P. Durbin (Ed.), *Research in philosophy and technology* (pp. 103–129). Greenwich: JAI Press.

Sintonen, M. (1990). Basic and applied sciences – can the distinction (still) be drawn?". *Science Studies, 3*, 23–31.

Skolimowski, H. (1966). The structure of thinking in technology. *Technology and Culture, 7*, 371–383. (Reprinted in Rapp 1974, pp. 72–85).

Veugelers, R., et al. (2009). *Evaluation of the finnish national innovation system – policy report*. Helsinki: Ministry of Employment and the Economy.

Winner, L. (1977). *Autonomous technology*. Cambridge, MA: MIT Press.

Ilkka Niiniluoto has made his academic career at the University of Helsinki: Master of Science in applied mathematics (1968), Doctor of Philosophy (1974), Associate Professor of Mathematics (Logic and Foundations) (1973–1977), and Professor of Theoretical Philosophy (1977–2014). He has been the President of the Philosophical Society of Finland in 1975–2015 and the editor-in-chief of *Acta Philosophica Fennica* since 1980. Niiniluoto has worked in philosophical logic, philosophy of science, epistemology, philosophy of technology, and philosophy of culture. His main works include *Is Science Progressive?* (D. Reidel, 1984), *Truthlikeness* (D. Reidel, 1987), and *Critical Scientific Realism* (Oxford University Press, 1999, paperback 2002).

Chapter 7
Changing Perspectives: The Technological Turn in the Philosophies of Science and Technology

Alfred Nordmann

Abstract The philosophy of science and the philosophy of technology share the same fate. The experimental turn in philosophy of science and the empirical turn in philosophy of technology open the black boxes of explanatory models and technical systems, and consider the creation of phenomena and artefacts. And yet, technology is viewed through the lens of science, subservient to or derivative of representation and the relation of mind and world. The philosophy of technoscience and an epistemology of working knowledge introduce a technological turn that affords a view of research as technological practice, both in science and in engineering.

Keywords Philosophy of technology • Empirical turn • Technoscience • Experimental turn • Technological turn • Working knowledge • Models and modelling

7.1 The Challenge

In the 1960s the philosophy of science was transformed in its encounter with the history of science, resulting in a collaborative venture by the name of "history and philosophy of science" (HPS). From here on, philosophy of science worked with case studies on an increasingly regular basis to reconstruct certain episodes from the history of science. The so-called "experimental turn" of the 1980s owed much to this interplay between philosophy and history. And although its guiding question remained quite traditional – namely, "How do the sciences achieve an agreement of representation and reality?" – the answers put forward have tended to break with tradition by focusing not so much on theory but on the role of instruments and experiments.

A. Nordmann (✉)
Technische Universität Darmstadt, Darmstadt, Germany
e-mail: nordmann@phil.tu-darmstadt.de

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_7

There is a flipside to the experimental turn, however, and it also needs to be told. For 30 years, if not more, philosophers and historians of science studied scientific practice and considered the role of instruments and experiments in laboratory science, and yet they disregarded for the most part the technological character of science and the relation between making and knowing. They looked at technological practice in scientific research but looked at it through the lens of a traditional philosophy of science with its preconception of science as an essentially intellectual endeavor and its attendant epistemological or metaphysical issues. Even more strangely, perhaps, the same can be said about some of the attempts to get a better understanding of engineering practice and technological artefacts which was a main concern of those who promoted various forms of an 'empirical turn' in science and technology studies and the philosophy of technology. It is in this sense that one might pose the question whether the *technological* turn even of the philosophy of technology is still outstanding: despite their empirical, experimental, and practice turns, is there a failure of most philosophy of science and much philosophy of technology to fully acknowledge the technological aspect of research? This seemingly paradoxical question rests on the admittedly contentious claim that there are different ways of taking technology seriously, some more cognizant than others of *homo faber* and the technological condition. This contribution seeks to substantiate this claim, sketchily in regard to the philosophy of technology, more extensively in regard to the philosophy of science.[1]

As for the empirical turn in the philosophy of technology, when opening the black box of technology, what was discovered for the most part was scientific research practice. This is evident for a book like *The Social Construction of Technological Systems* (Bijker et al. 1987). The "social constructivism" in the title originated among competing ways of accounting for the adoption of scientific beliefs and was transposed to the different way of accounting for the adoption of technologies. It offers a framework for studying the deliberation and settlement of the question whether some new technological proposition represents progress. Similarly, close attention to the methods and reasoning processes in engineering and the technological sciences typically revolves around rather intellectual conceptions of research. For example, technological research is said to concern the truth or falsity, the validation and acceptance of claims regarding the proper means towards

---

[1] This claim is emboldened by Paul Forman's (2007) reflection on "The Primacy of Science in Modernity, of Technology in Postmodernity, and of Ideology in the History of Technology." Forman argues that most philosophers and historians of technology failed to notice, take seriously, or critique the primacy of technology but discuss technology as subservient to the progress of science and society, that is as a means for the advancement of non-technical goals. My claim finds support also in recent remarks by Martina Heßler who argues that technology and humans are missing in contemporary studies of "technology in context." According to Heßler, these situate black-boxed technological systems and devices in a social context that consists of actors who shape or use technology. She calls for a "historische Technikanthropologie" which considers how humans understand themselves in relation to technology (Heßler 2013). Heßler's anthropology of technology cannot be undertaken if one presupposes the primacy of science and society, if one privileges the head over the hand, if one is preoccupied with the relation of mind and world. See also Nordmann (2015a).

some end (Kornwachs 2012). In a different vein, the influential research programme on the Dual Nature of Technical Artifacts was motivated by the theoretical question of how to conceive the causality of human intentions in the sphere of physical materiality (Kroes and Meijers 2006). Here, technical artifacts are objects of knowledge primarily in that they are subject to causal processes which need to be appreciated and understood. The empirical turn of the philosophy of technology was thereby modelled on empirical science. Finally, recent developments in philosophy of science – be it in regard to the experimental turn, in regard to explanatory mechanisms, or in regard to models as mediators – were adapted to describe engineering research.[2] Again, this was done not in order to exhibit and explore strategies of technological reasoning in the sciences, but instead to highlight that engineering research is quite like that of science under the best description of how science manages to forge agreement between theory and reality or mind and world.

The philosophy of technology asks entirely different questions in that it begins right in the middle of things with the fact that there is technology. It does not consider the meeting of mind and world as a precarious encounter, haunted by the possibility of skepticism and requiring cumbersome reconstructions of how the agreement between theory and reality can be ascertained. When things are constructed, built or made, human thinking and physical materiality are inseparably intertwined from the start – as inseparable as the head and the hand and the eyes and the world. Accordingly, if scientific or engineering research was to be understood in terms of technology – rather than technology in terms of scientific or engineering research – the researchers would have to be treated as builders and makers who use theories, algorithms, models, and material objects as their tools and materials. And opening the black boxes of technology, one would not find the underdetermined decisions that were made by scientists, engineers, and developers. Instead, one would find the accomplishment of a working order, that is, an arrangement of people and things such that they manage to work together in a socio-technical system. This way of looking at scientific and engineering research opens up new avenues for an epistemology of so-called technoscience, that is, for the knowledge production of *homo faber* who uses scientific theories as tools for the acquisition of demonstrable knowledge of how things can work together. In order to expose these avenues of philosophical investigation, it helps by way of background to contrast in broad strokes how the philosophy of science and the philosophy of technoscience take technology seriously.

7.2 Competing Projects

Starting with Immanuel Kant, the fact of science mattered to philosophers because it shifts the question from "is it possible to acquire true knowledge of the external world?" to "how is such knowledge possible since, evidently, it exists?" Seeking an

---

[2] Especially in the work of Mieke Boon, e.g. Boon (2012).

answer to the latter question, philosophers and historians of science had many theories to debate – various kinds of realism and various kinds of constructivism can account for this possibility. And in the course of debating these metaphysical theories, philosophers and historians of science learned to consider scientific knowledge production in terms of practice, that is, conceptual as well as laboratory practice.

In recent years, the competing project of history and philosophy of technoscience has taken shape – a project which, to a certain extent, turns previous philosophy of science not so much on its feet as on its hands by understanding science as technology. It presents itself as a complementary project that does not seek to displace received philosophy of science. And yet it articulates a concept of research that unsettles the normative conceptions of science of the nineteenth and twentieth centuries and which the traditional philosophy of science must, of necessity, reject. At stake here is nothing less than the relationship between science and Enlightenment, for instance the idea that there is a critical point in the quest for truth at which our ambition to control and design the world reaches its limit. The philosophy of technoscience is based on the notion that all science is technology and needs to be considered from the perspective of philosophy of technology. Put in quite general terms, it does not revolve around the metaphysical question whether or how there can be agreement of theory and reality. Since the technosciences are not concerned primarily to represent reality, the competing project's main question addresses the relationship between knowledge and skill in building and making, in manipulating and modeling.[3] In this context, traditional concepts such as explanation and understanding, validation and objectivity need to be understood in the way they apply to technical construction procedures or the technoscientific knowledge of how things can work together in an apparatus, a device, an experimental or technological system. And although such knowledge often serves to advance social progress, as a rule we do not imagine that it entails the Enlightenment ideal of speaking truth to power.

Evidently, then, these are two ways of taking technology seriously. The first is to consider the role of technology in scientific practice or to open the black box and study how scientific practice inhabits technical systems. The second is to adopt a vantage point that is not in the sphere of opinion or belief but in the sphere of working knowledge of how things work. In the following, the transition from the former to the latter will be rehearsed in four steps: I begin with the experimental turn initiated by Ian Hacking which, despite elucidating the relationship between intervention and representation, in no way calls into question the metaphysical preoccupation of the philosophy of science. After developing further the difference between science and technoscience, I will show in the third step that various approaches in today's philosophy of science undermine its propositional preconception of science (even if they do not question it) by making increasing use of a technical idiom to reconstruct the agreement between representation and represented reality. It is only then, in the fourth step, that the new questions of the philosophy of technoscience can be appreciated. They become discernible through a change of perspective when

---

[3] If metaphysics queries the preconditions of knowledge as agreement between mind and world, one might speak of metachemistry as exploring the preconditions of knowledge through making and building; see Nordmann (2013).

the technical idiom of recent philosophy of science gives way to the technological turn that yields the philosophy of technoscience.[4]

7.3 Hacking's Overture

One of the most influential books in philosophy of science of the last 30 years is Ian Hacking's *Representing and Intervening*. This book problematized the guiding question of philosophy of science; it took technological practices seriously from the point of view of the philosophy of science and established a new strand of investigation known as the philosophy of scientific experimentation. Despite all this and despite the book's plea for a fundamental reform of philosophy of science, however, ultimately it continues the classical project of philosophy of science, its guiding question included. No matter how much it emphasizes the technical intervention of the experimenter, the book sees in technology nothing more than instrumental means that are a subservient to the achievement of an intellectual goal.

This is how Hacking identifies the philosophical problem that he seeks to address (1983, p. 130): "By attending only to knowledge as representation of nature, we wonder how we can ever escape from representations and hook-up with the world." This worry reflects the traditional concern of the philosophy of science that can be traced back at least to Kant: Given that the human mind produces representations only, how can one be sure that these actually agree with an externally given world? To be sure, Hacking's formulation suggests that there might be something wrong with this question. But as we shall see, he critiques only a specific interpretation of this question and leaves its general version intact.

Hacking inserted a "break" between the main parts of his book that deal with representation and intervention respectively. Here he cautions the reader against assuming that representation and intervention are alternative, even mutually exclusive practices. He points out that he is not arguing against philosophy's predominant interest in representation but only against the view that representation is always conceptual or theoretical. Thus he critiques the so-called spectator theory of knowledge, not least because it fails to acknowledge that representations are *made*, that they are generated – not least by means of intervention. This fault of the spectator theory becomes especially apparent when the relationship between representation and the represented is reduced to a relationship between theory and world (ibid.):

> Incommensurability, transcendental nominalism, surrogates for truth, and styles of reasoning are the jargon of philosophers. They arise from contemplating the connection between theory and the world. All lead to an idealist cul-de-sac. None invites a healthy sense of reality. […] In our century John Dewey has spoken sardonically of a spectator theory of knowledge that has obsessed Western philosophy. If we are mere spectators at the theatre of life, how shall we ever know, on grounds internal to the passing show, what is mere representation by the actors, and what is the real thing?

---

[4] Sections 7.3, 7.4, 7.5, and 7.6 have been translated and adapted with the help of Kathleen Cross from Nordmann (2012a).

Dewey and Hacking were not the only ones to mock western philosophy's obsession with how the knowing subject can get a grip on the world after it has taken its seat in the spectators' auditorium. Cultural critics Martin Heidegger and Hannah Arendt also exposed this obsession as a fundamental error of modernity: first the individual withdraws into the confined space of his or her perceptual apparatus and, ultimately, into his or her skull, in order then to ask how there can be such a thing as agreement between inside and outside, between mind and world, between theory and reality. These questions about how such agreement is possible, how it is generated and how it can be validated, would not even arise without the initial withdrawal. René Descartes demonstrated this especially clearly. His philosophy and his science begin expressly with the somewhat absurd effort of settling in front of the fireplace and trying to banish all ordinary commitments and empirical matters from his thoughts until nothing remains but a thinking thing, the pure spectator.[5] His philosophical thought experiment conjured up the image of a human being overwhelmed by the chaotic multiplicity of sense perceptions, desperate to find an Archimedean point from where order can be brought into chaos and the world can be rendered comprehensible. Scientific research appears accordingly as the production and interpretation of data: observation and experiment confront scientists with impressions and measurements that are brought together using a theory or a model and that, where possible, include predictions that allow for the agreement of theory and reality to be tested.

Although no one will doubt that Descartes' thought experiment has proven extraordinarily fruitful for science and philosophy alike, it is not difficult to understand just how artificial it is to suppose the opposition between the real and the conceptual, between the world to be known and the knowing subject, between raw data and their interpretation. From the point of view of the philosophy of technology there are two reasons why this opposition cannot serve as a point of departure.

First of all, there may have been a time when natural phenomena and sense impressions constituted a bewildering multitude and could be ordered by intellectual means alone. In the meantime, however, several thousands of years of scientific and technological history have gone by so that people today live in a largely disenchanted world – a world that is intellectualized, predictable and technological, one that provides structure for largely routinized action. Nowadays it takes a great deal of technical effort even to produce a bewildering diversity of phenomena that require interpretation, as, for example, in scientific experiments in which unprecedented amounts of data are generated. Yet to the extent that these phenomena are generated using the methods and means of a science-based technology and are recorded and interpreted using technical media, it is more than strange to postulate that the goal of knowledge is to bridge two spheres of mind and world, theory and reality, imagined to be logically independent of one another. Instead, it would be more apt to speak of two mutually calibrated kinds of apparatus encountering

---

[5] Descartes, *Discourse on Method* (1979) and *Meditations on First Philosophy* (1996).

one another – namely, on the one hand a scientifically and technically defined experimental system, including the data sets generated from it, and on the other hand a scientifically and technically defined system of thinking and perceiving, including observational instruments and representational techniques. And indeed, this is how Heidegger and Arendt reconstruct the beginnings of modern science with Galileo and Newton, and it is how they reconstruct the history of epistemology from Descartes through Kant to Heisenberg. Both of them cite Heisenberg's statement that wherever humans look, they always encounter themselves. Whereas Arendt (1958, p. 261) is thinking here of Galileo who located the Archimedean point within the knowing subject and its techniques of observation and measurement, Heidegger (1967, p. 332; 1977, pp. 93 and 89) traces it back to the construction of a continuous nature in Newton's mechanics that serves to mathematically represent the succession of events.

There is a second reason why, from the point of view of technology and of the philosophy of technology, it is not plausible to position research that builds theory, interprets data, and explains things opposite a mind-independent reality. This is because scientific research is not exclusively or predominantly concerned with producing theories, hypotheses and models, that is, with linguistic statements that are then tested as to their truth or falseness or their empirical adequacy. If one wants to make a thing work, one doesn't first make a truth claim; rather, one sets up a technological system that is judged by its performance, according to what it can do or what it affords. To be sure, there are technological systems that presuppose a correct representation of reality – philosopher of science Michael Friedman (2010) would probably mention the calendar at this point, which is based on precise astronomical knowledge. These are not the rule, however, and even clockwork represents nothing – or, at most, it represents itself as a specific kind of working order that "keeps time" in the sense of exhibiting regularized motion. So, if the opposition of mind and world or of theory and reality makes sense when hypotheses are formulated and tested, it does not make sense for other scientific activities: at times, theoretical statements are formulated, at other times things are made to work, at yet other times phenomena and processes are experimentally generated and stabilized, "proofs of concept" are established, interesting properties of objects are demonstrated and skills for their demonstration acquired. Leafing through scientific journals one still finds, of course, articles in which hypotheses are tested by means of newly acquired evidence; what one finds above all, however, are articles in which the controlled growth of carbon nanotubes is produced or the pharmaceutical efficacy of a substance established, in which an improved method of physical intervention or of visualization is demonstrated, in which a model system is put into operation. This increase in technical capabilities can also be described as growth of objective knowledge and is presented in just these terms in journal publications. But by what criteria can we speak of knowledge here? This is a notoriously difficult question for philosophy. At any rate, we won't get very far if we interpret this knowledge as an agreement between mind and world – not least because mind and world cannot be distinguished in a clockwork, a computer simulation, a genetically modified lab mouse or the demonstration of a causal mechanism (Nordmann 2006, 2012b).[6]

From the point of view of the philosophy of technology it is easy to appreciate Dewey's and Hacking's surprise at a spectator theory of knowledge and its artificial conception of the basic epistemological or metaphysical problem. We might suppose that it is at this point that Hacking's transition from *representing* to *intervening* occurs, and that he will now turn away from knowledge as a representation of nature and turn instead towards a technological concept of knowledge in doing and making. Hacking proceeds differently, however, arguing merely against the implied passivity of the spectator, against the division between thinking and acting and against the "theoretical" interest of philosophy of science that is related only to thinking (1983, pp. 130–131 and 160; *cf.* van Fraassen 2008).

> Yet I do not think that the idea of knowledge as representation of the world is in itself the source of that evil. The harm comes from a single-minded obsession with representation and thinking and theory, at the expense of intervention and action and experiment. (pp. 130–131.)

Thus Hacking does not end up questioning the supposition that scientific knowledge consists in representing and depicting nature; his critique is focused on the notion that the representation and depiction of nature is only a matter of theory, of thinking and not also of doing and making. His book is not interested in *homo faber* but only in *homo depictor*, as he puts it a few pages farther on. When this *homo depictor* makes an image in his cave or carves a figure from wood, he is always already in the business of creating likenesses, with each artefact formulating the "characteristically human, thought […] that this wooden carving shows something real about what it represents" (p. 136; cf. pp. 132, 137f.). With his attention focused on experimental intervention and the creation of phenomena, Hacking eliminates the spectator and includes technical action within philosophy of science – but he does so on the presupposition that every presentation of a symbol or phenomenon, even the technical stabilization in the laboratory of a physical process implies a claim about the truth or correctness of a representation, namely, that this here represents something real.[7] According to Hacking, then, the main problem and guiding

---

[6] Since the time of Plato, Western philosophy has worked with the definition of knowledge as "justified true belief". Someone who demonstrates they have acquired a skill is not formulating a belief, that is, a proposition that could prove either to be true or false, justified or unjustified. To put it differently again: anyone who draws on the intellect to establish agreement between mind and world has to worry constantly whether it is even possible to conform to mind-independent reality. This is tantamount to the worry that the presumably discovered token of reality might be nothing but an artifact of the chosen procedure or method (here, "artifact" has a negative connotation). Someone interested in producing a therapeutic active substance, on the other hand, is seeking to create an artifact (positively connoted). Accordingly, they will not seek to sort out whether this artifact owes its existence to mind-independent reality or to human intervention.

[7] From the perspective of the philosophy of technology Hacking's anthropological "fancy about the origin of language" (idem, p. 135) appears just as peculiar as the spectator theory of knowledge. After all, cave paintings and wood carvings are significant and contribute towards the organization of social life even without serving predominantly as claims about what is real and without triggering

question of science and philosophy of science remains that of the agreement between a representation and the thing represented. Hacking's experimental turn sees the practical solution to this problem in instrumental realism and thus in the fact that the realism of scientific representations owes simply to the fact of their experimental realization, that is, to the technical stabilization of phenomena or the active engendering of effects that goes into the making of these representations. To the extent that this instrumental realism serves to solve the problem of realism, scientific instruments and experimental technology are merely the means by which science captures reality. Thus Hacking does not render research as technological activity and does not situate science in its technological setting. When phenomena are created or generated in the laboratory, as far as Hacking is concerned, this means principally that someone has succeeded in creating the likeness of law-like nature in the laboratory – a technically reproducible phenomenon represents the laws to which it owes its own existence (idem, pp. 222–224).

Finally, should any further evidence be required that Hacking is still addressing the guiding question regarding the agreement between representation and reality and merely answering it in a new way, then it is the historical circumstance that his project was indebted to the debate about realism and constructivism. Indeed, the first sentence of Hacking's book is: "Rationality and realism are the two main topics of today's philosophers of science."[8] The dispute about realism and constructivism is definitively of great concern to *homo depictor*, along with the issues of physicalism and the unity of the sciences, incommensurability and scientific progress, internal and external influences, and rational reconstructions of theoretical dynamics. All of these issues were hotly debate in 1983 and for quite some time afterwards. It is worth asking, however, what has become of them in recent years. Which philosophers or historians are still interested today in theory choice and its rational reconstruction? Does anyone still get excited about the question of whether biology is reducible to physics? Is anyone driven to defend realism against constructivism?

---

debates of the kind "No, not that, *this* here is real". They might serve ritualistic functions, for example, by instituting rhythms and patterns of action or dependency. Putting the point in a Wittgensteinian manner: A wood carving cannot say that something does not exist or is falsely represented, and thus it cannot say that something exists and is correctly represented. In a discussion of two wood carvings it might well be said under very special conditions that the one is "more similar" than the other (assuming both carvings make the same "claim" and are trying to be representations of the same object), but even then there would not be any criterion for telling when a carving "says" something that is "correct" or "false" about reality. The fact that Hacking presents this archaic representational practice as a prototype for the formulation of propositional theories reveals that he is thinking and writing in the context of the traditional debate of scientific realism. By positing that, when a wooden figure is carved, a thought is formed, he has not distanced himself from a philosophy that is in love with theory and thought. (On this issue, Ludwig Wittgenstein takes a more subversive and fruitful stance: His picture theory of language is based on the discovery that under quite specific circumstances a game with model cars that initially represents nothing can be transformed such that the toys are now said to stand for actual cars in an accident. Only now and only thus does playing with cars become modeling and representation of reality, see Wittgenstein (1994, p. 279).

[8] This sentence opens the "analytical table of contents" in Hacking (1983, p. x).

Is anyone out to prove the real reality of theoretical entities? What are, then, the key questions of today's philosophers of science? This leads us to the situation of philosophy of science 30 years after Hacking's intervention.

7.4 Science and Technoscience

This detailed discussion of Ian Hacking's *Representing and Intervening* illustrates that a great deal more can be done once one starts considering the interplay between the technical creation of phenomena and their theoretical representation. Indeed, philosophers of science have since developed a rich vocabulary for describing in technical terms how the agreement of representation and represented is forged. As in Hacking's account, these descriptions paint a technological picture of theoretical practice, but barely do justice to technoscientific knowledge production and have nothing as yet to say about the increase of objective knowledge through the accomplishment in a laboratory of the controlled growth of carbon nanotubes. Once we are interested not in the subservient role played by technology for the agreement of representation and reality but instead in the advancement of knowledge through the technical manipulation of complex phenomena and processes, we need a philosophy of science that is informed by philosophy of technology.[9]

In order to illustrate all of this, let us take a somewhat closer look at how science and technoscience have been contrasted – if only implicitly – so far. This contrast provides the background against which some recent developments in philosophy of science can be recognized and understood – developments that have pushed well beyond the bounds of classical issues of rationality and realism and have infused a technical idiom. By employing this idiom philosophers of science are already proceeding in a technoscientific manner. And yet, they are not reflecting upon technoscientific research but still describe practices of building, employing, and testing theory.

The sciences are concerned with the theoretical representation of a mind-independent reality, that is, they need to vouchsafe that theory and reality are meaningfully related and yet independent of one another. To this day, physics does this in an exemplary manner and has therefore been considered to be the paragon of science. The search for the Higgs Boson might serve as an example. The Large Hadron Collider that was devised for its detection is a gigantic technical set-up. And yet, whether or not the Higgs Boson exists is definitely not a technoscientific question but an utterly classic case of testing a hypothesis or a theory. The question of the

---

[9] What does the philosophy of technology have to offer the philosophy of science apart from a different point of departure? This requires a more informed answer than I am able to offer in this context, but one which issues in a challenge to the philosophy of technology regarding, for example, the epistemology of working knowledge. The attentive reader will notice, at any rate, that the present essay doesn't have much more on offer as yet than Heidegger's technological understanding of modern science.

existence or not of the Higgs Boson is tied to the truth or falsity or empirical adequacy of the best available theoretical description of reality. And precisely because so much technical effort is invested in the detection of the Higgs Boson, the credibility of the research hinges crucially upon our ability to defuse the suspicion that the evidence for the Higgs Boson is a technical artifact. A great deal of methodology and reasoning goes into the argument that the Higgs Boson is something that exists objectively. The proper and significant relationship between theory and reality is such that the evidence is not caused by our research technologies but only brought to light by them.[10] Science stands in the tradition of Enlightenment with the search for the Higgs Boson and for evidence that serves the theoretical purpose of critiquing or validating knowledge claims.

In contrast to this, the technosciences are dedicated to the acquisition and demonstration of capabilities of control. This includes fundamental capabilities of modeling and predicting, of visualization and manipulation. When it is demonstrated, for example, that the growth of carbon nanotubes can be controlled or that the replication of a cancer cell is halted by a certain active substance, or that the error rates of a type of detector on the Large Hadron Collider can be calculated and corrected, there does not arise the question of carefully holding apart what exists naturally and what owes to technical intervention. Rather, doing something by means of technology entails having the necessary knowledge to influence systematically a complex set-up that is as “natural” as it is “human-built.”[11] Thus technoscientific research is defined by a technological mode of knowledge production and by its specific interest not in the nature but in the potential or power of its objects of research. It would be misleading, therefore, to define the sciences as fundamental and intellectually curious while conceiving the technosciences as application-oriented and utility-driven. Indeed, there is basic technoscientific research that involves the acquisition and demonstration of fundamental capabilities of technical access, manipulation, and control; these capabilities often entail further development of the research technology and only occasionally feed into other technological contexts where things are made – the so-called applications. Materials research and nanotechnology, computer science and synthetic biology are the paragons of technoscience, although ultimately all the sciences – including the human and social sciences – take on the

---

[10] This kind of ontological concern, even anxiety, regarding the status of evidence for theoretical knowledge of reality contrasts with the ontological indifference of tinkerers, engineers or technoscientists. On “ontological indifference” see Galison (2016) and (though they do not refer explicitly to “ontological indifference”) Daston and Galison (2007, Chap. 7, especially pp. 393 and 414).

[11] Peter Galison’s (1997) had previously shown that although experimental systems such as the Large Hadron Collider are completely geared towards testing a theory and are thus classically scientific in nature, only a fraction of the research that is done there is scientific in this sense of the term: a majority of the work is dedicated to developing and modeling detectors, that is, it is concerned in a technoscientific sense with acquiring and demonstrating fundamental capabilities of observational control.

character of technoscience when they create or modify technical settings and, for example, capabilities of facilitation, deliberation, and cooperation.[12]

This rather superficial story of Enlightenment science and the recent acknowledgment of technoscience helps us understand some current trends in the philosophy of science. First, philosophical accounts of the ways in which scientists forge agreement between theory and reality increasingly proceed in a technological idiom and employ the vocabulary of engineering. And second, although science is not explicitly regarded as technology, there are hints in the direction of technological accounts of research practice – though, to be sure, one cannot speak as yet of an established philosophy of technoscience.

7.5 Philosophy of Science in the Idiom of Technology

The first observation can be underpinned by referring to just three current approaches that describe scientific praxis in a technical idiom. To begin with, there are Margaret Morrison's (1999) and Nancy Cartwright's (1999) reconstructions of scientific modeling. Foremost in these accounts are notions of *fitting* and *tuning*: models serve to produce a fit between theories, principles or concepts on the one hand and phenomena on the other. In these local processes of calibration the model functions as a working tool, as medium and mediator for coupling special conditions to general principles. While fitting and tuning are technical procedures, they serve the theoretical purpose of achieving (at least local) agreement between representation and reality, which places this discussion in the familiar context of the philosophy of science.[13]

Another influential and stimulating proposal is the reconstruction of scientific explanation as the specification of a mechanism that generates a phenomenon or a process.[14] In contrast to the philosophical tradition this proposal speaks of "explanation" without reference to theories, models, generalizability or the logical relations between propositions. If this conception still relates to the classical question regarding the agreement of representation and what is represented, this is because it seeks to exhibit a real structure or a real mechanism and to do so in a clearly comprehensible,

---

[12] The distinction between science and technoscience has been presented in greater detail and with more sophistication in (Bensaude-Vincent et al. 2011). This text also discusses methodological issues for the study of science and technoscience and how to conceive their relationship.

[13] This motif can be traced by to Wittgenstein (1993). When combined with a different concept of the model and an iterative method borrowed from software engineering, fitting and tuning can also serve to generate not so much a representation that accords with reality but rather a technical model-system that is a substitute for reality. The distinction between representational *models of* and substitutional *models for* cannot be explicated here but its exploration is one of the tasks of philosophy of technoscience; cf., for example, Lenhard (2011).

[14] The classic source for this view is (Machamer et al. 2000). The technical significance of the mechanisms discussed here comes from Glennan (1992).

intellectually tractable manner. We are supposed to see that the really observed events happen according to the mechanism that has been identified.

A third instance of a technical vocabulary taking hold concerns "robustness" as a criterion for the "reliability" of statements.[15] Robustness typically signifies that theories do not have to be true, truthlike or empirically adequate; they do not have to agree with reality, make precise predictions or elicit reproducible confirmations. Instead, they are required to prove themselves like rules of thumb in contexts of use. To be sure, as a criterion robustness often serves as a measure of the truth or acceptability of theories as representations of reality, though robustness *par excellence* can serve to validate, entrench, institute a non-propositional material system or practice.[16]

All three approaches explicate the basic supposition that the aim of science is to represent phenomena and processes. To this extent their respective authors are making use of a technical vocabulary merely as a practical aid for characterizing something that is primarily theoretical or intellectual. And yet the scientists described in these terms no longer appear as critical minds and Enlightenment thinkers who advance overarching historical developments such as theoretical unification or physical reductionism, such as intellectualization or rationalization. Instead, the scientists so described are delivering piecemeal solutions to locally specific problems. In these descriptions of scientific practice, so much emphasis is placed on the technical requirements for the forging of agreement that the ideals of knowledge and science on which this is predicated recede into the background. Once the great questions of rationality and realism, of intellectual progress and true knowledge of the world, have disappeared from the scene, what remains is an interplay between representation and intervention so intimate that the distinction between science and technoscience appears to be redundant. Many philosophers of science gladly embrace this appearance, since it suggests that they are already attending to all aspects of knowledge production, and that they need not engage with radically different epochal or disciplinary conceptions of knowledge-producing research. Accordingly, they quickly convince themselves that with a diffuse concept of scientific practice one is already achieving all there is to achieve.[17] It is to counter this

---

[15] Cf., among others, Roush (2005).

[16] This list makes no claim to be exhaustive. For example, Bayesian philosophy of science has also adapted the idiom of engineering in that it develops instruments for tracking changing degrees of belief.

[17] Compare Nordmann (2015a). Instead of relying on such a diffuse concept of practice, Martin Carrier pursues another strategy that allows him to ignore the fundamental difference between scientific and technoscientific knowledge production. Carrier refers to two kinds of continuity. One is the continuity of science's orientation toward utility, which underpins the entire development of modern science. The other is a methodological continuity in terms of causal analysis, modeling and validation. On this view, the control of phenomena and the descriptions of the world are always in interaction with one another. The appearance in the eighteenth to twentieth centuries of a seemingly purely theoretical science should not distract our attention from this continuity; see his (2011). In my view, however, Carrier underestimates (1) the powerful influence of a notion of pure science – however elusive it may be – upon a philosophy of science that has in the 19th and most of the twentieth century articulated and valorized a very specific understanding of what science

somewhat self-satisfied obliviousness that it is worthwhile to differentiate between science and technoscience. On the one hand we have the myth of a hypothesis-testing science that seeks Enlightenment and produces true or empirically sound knowledge about reality; this myth continues to exert a latent influence and therefore needs to be acknowledged even where one otherwise shies back from speaking of "science" in the singular and deals instead with the diversity of special sciences. On the other hand we have the equally longstanding and influential (albeit not in the philosophy of science) myth of a technoscience that pursues innovation and acquires and demonstrates capabilities of control. Indeed, articulating the technoscientific ideal of research also helps us understand the classical concept of science and the values with which it is associated.[18]

7.6 Beginnings of a Philosophy of Technoscience

Even where it adopts a technical vocabulary and thereby subverts the image of the scientist as advocate of Enlightenment, today's philosophy of science is interested in the relation between mind and world, between theory and reality. It thereby views science through the lens of the history of ideas. Though it is difficult in most fields of research to describe scientific progress as the advance of powerful ideas or theoretical concepts, ideas set the terms when one establishes a local fit between theory, its models, and the phenomena, when one specifies a mechanism that articulates these models, when one probes the robustness of some theory as a tool for describing specific bits of reality. And so the philosophy of science of the last 30 years or so has acknowledged the technical requirements for representing reality and simultaneously moved away away from questions associated with Kuhn and Lakatos, Nagel and Putnam, Harman and Kitcher. Rational reconstruction and the problem of theory choice, physicalism and reductionism, theoretical unification and inference to the best explanation haunt current debates in philosophy journals like ghosts from the past.

The concurrent trends of waxing and waning philosophical interests suggest that there is no longer any unquestioned, unconditional belief in overarching trajectories of scientific development or the progressive advance of a scientific worldview such

---

believes itself to be; (2) the difference between controlling phenomena and explaining the world with regard to notions like "causal analysis," "modeling," or "validation" – these terms (like "knowledge," "theory," "explanation," etc.) have different meanings in scientific and technoscientific research; (3) the emergence of new methods that unsettle or interrupt the continuity – such as iterative procedures for increasing the complexity of models or explanatory inferences to an underlying dynamic from the similarity between two visualizations. Conversely, I am happy to admit that I overestimate discontinuity and, above all, that I fail to dwell on the fact that, of course, the radically different modes of knowledge-production are actually in conversation with one another, provide correctives to and complement one another.

[18] This is Bernadette Bensaude-Vincent's (2009) view of what a philosophy of technoscience ought to do.

as the *Wissenschaftliche Weltanschauung* of the Logical Empiricists. The shift of attention to particular, mostly local ways of forging agreement between theories, models, and bits of reality has its flipside in skepticism regarding universalizing stories about the unending quest for knowledge, about the Enlightenment project with its emphasis on the critical attitude of science, about the emergence of a physicalist account that overcomes disciplinary divisions. And this is where, historically speaking, the philosophy of technoscience enters in. From this shift of attitude, it is but a small step to a complete change of perspective that sees the process of research not as a theoretical representation of the world but as the piecemeal technical appropriation of phenomena – and thereby dispenses with the preconception that the latter is always subsidiary to the former. This preconception and thus the original conceit of the philosophy of science has been undermined by the philosophers of science themselves. It is now only a ghost or vestige of the grand universalizing stories that for the most part have already been abandoned. By assuming that research consists, roughly speaking, in the piecemeal technical appropriation of phenomena, we would consider the existing repertoire of scientific theories as part of the toolbox on which researchers can draw in order to demonstrate the ways in which things work together within a technological working order. The scientific development of theories, hypothesis testing and the evaluation of truth claims would thus appear as subservient to the production of technoscientific knowledge. Where, however, is this change of perspective towards a philosophy of technoscience taking place? Here, too, I will name just a few developments from the philosophy of science that are actually indicative of a philosophy of technoscience.

First and foremost there is the philosophy of computer simulation as pursued by Paul Humphreys (2004) and Johannes Lenhard (2011; Lenhard et al. 2007), for example (see also Winsberg 2010). They are looking at a research technology that fundamentally calls into question received concepts of philosophy of science, or rather reconfigures them: what is an experiment, what is a model, what is an explanation, what does it mean to understand something, what is the relation between making and knowing? The answers to these questions pertain also to biomedical research with model organisms (Fox-Keller 2000) and to the "real-world simulations" explored by Astrid Schwarz and Wolfgang Krohn (2011). The received view that models are models of theories and that they offer an intellectually tractable means of conceptualizing mind-independent reality cannot be upheld for these animal and computer models. These are material systems that stand in or substitute for reality to the extent that they gained real-world complexity through the iterative methods of their construction. Rather than become intellectually tractable through the reduction of complexity, they are meant to absorb as much as knowledge of reality as possible, such as to afford experiments on reality by way of experiments on the models (Nordmann 2006).

If this is the most prominent technoscientific topic, there are various other themes and approaches that point in the direction of a philosophy of technoscience and deserve mention. Growing out of the concern with computer modelling there is increasing interest in data-intensive science and data curation.[19] Other contributions include – in no particular order and by way of example only – those by Hasok Chang (2007), Eran Tal (2016), Kenneth Waters (2008), Maureen O'Malley (2011), Robert Batterman (2009), or Anna-Sophie Heinemann (2013). Chang and Tal look at metrology and its iterative procedures for internal validation. Iterative procedures are found in software engineering, metrology, synthetic biology and climate modeling. They allow for technical systems to validate themselves by way of being optimized in respect to their desired performance, rather than being validated through brute confrontation with a logically independent external reality. According to canonical views in the philosophy of science procedures of internal validation are threatened by circularity and by presuppositions of what it is to be shown. The task of a philosophy of technoscience, however, would be to explore the grounds for justifying such procedures – not least when their practitioners refer to the fact that they work technologically: How are we to understand, for example, that the mere demonstration and description of a thing constitutes a substantial proof of something?[20]

The reference to a mere thing highlights another recent development that might herald a philosophy of technoscience. In the Kantian tradition it became a commonplace to deny that things could be an object of knowledge – for philosophers like Ernst Cassirer, Ludwig Schlick, Ludwig Wittgenstein the world of modern science is a world of facts and not of things. Recent years have witnessed a renaissance of things in technoscientific research practice and in philosophical discussions.[21] The concern here is not with sense data, functions, relations, substances and properties, or facts of experience, but with things, their power and potential, what they afford, the surprises they hold in store, their capabilities of doing work in a working order of things (Baird and Nordmann 1994; Baird 2004; Nordmann 2015b).

7.7 Outlook

After this superficial and admittedly selective inventory of current philosophy of science, a new understanding of research practice and its cultural significance, one that draws upon philosophy of technology, is already becoming apparent – even if it is made explicit only rarely. Now, the same task presents itself for a philosophy of technology that has taken an empirical turn but still needs to be grounded in working knowledge of the working order of things. This would be a philosophy of technology that does not begin with the engineer who solves a problem, nor with a user who takes up a device, nor a philosopher who marvels at the artefact. It would begin in Thomas Hughes's (2004) human-built world and with the affordances of

---

[19] See, for example, the work of Sabina Leonelli, such as her (2015).

[20] And thus it is a research topic for the philosophy of technoscience to analyze and clarify the notion of what is a "proof of concept."

[21] For an overview and particular investigations see (Bensaude-Vincent et al. 2016).

Rom Harré's (2003) apparatus-world complex, that is, with a world that is not primarily an object of knowledge but a product of knowledge. Rather than through deliberation from a critical distance we get to know this world through participation and the fact of being immersed in it – and we may find that the way things work in this world are not intellectually tractable but knowable and predictable simply because this world is a *verum factum*, that is, a world that we have made and that has made us.[22]

In the 1960s the philosophy of science was transformed in its encounter with the history of science, resulting in a collaborative venture by the name of "history and philosophy of science" (HPS). Nowadays and continuing on from Hacking's overture, HPS is transformed in its encounter with the philosophy of technology, resulting in a collaborative venture that might go by the name of "history and philosophy of technoscience" (HPtS). And just as Hacking's experimental turn harbored the momentous transformation that needs to be seen through by way of a technological turn, so the empirical turn of the philosophy of technology is complemented and consummated by its technological turn.

References

Arendt, H. (1958). *The human condition*. Chicago: University of Chicago Press.

Baird, D. (2004). *Thing knowledge*. Berkeley: University of California Press.

Baird, D., & Nordmann, A. (1994). Facts-well-put. *British Journal for the Philosophy of Science, 45*, 37–77.

Batterman, R. (2009). Idealization and modeling. *Synthese, 169*, 427–446.

Bensaude-Vincent, B. (2009). *Les Vertiges de la Technoscience: Façonner le Monde Atome par Atome*. Paris: La Découverte.

Bensaude-Vincent, B., Loeve, S., Nordmann, A., & Schwarz, A. (2011). Matters of interest: The objects of research in science and technoscience. *Journal for General Philosophy of Science, 42*, 365–383.

Bensaude-Vincent, B., Loeve, S., Nordmann, A., & Schwarz, A. (Eds.). (2016). *Research objects in their technological setting*. Abingdon: Routledge. forthcoming.

Bijker, W., Hughes, T., & Pinch, T. (Eds.). (1987). *The social construction of technological systems: New directions in the sociology and history of technology*. Cambridge, MA: MIT Press.

Boon, M. (2012). Scientific concepts in the engineering sciences: Epistemic tools for creating and intervening with phenomena. In U. Feest & F. Steinle (Eds.), *Scientific concepts and investigative practice* (pp. 219–243). Berlin: De Gruyter.

Carrier, M. (2011). 'Knowledge is power', or how to capture the relationship between science and technoscience. In A. Nordmann, H. Radder, & G. Schiemann (Eds.), *Science transformed? Debating claims of an epochal break* (pp. 43–53). Pittsburgh: Pittsburgh University Press.

---

[22] In a programmatic paper, this hint to Giambattista Vico cannot be articulated. We know the things that we have made in virtue of having made them – this general idea appears also in the works of Francis Bacon and plays a role in contemporary synthetic biology (there it is often attributed, somewhat misleadingly, to Richard Feynman). This knowledge is public and objective, thus different from skill or personal knowledge. It does not rely on correct representations of the things we have made. Before or after the making, such representations may well be possible, even necessary but, as such, different from the knowledge of how things work together in a working order.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.

Chang, H. (2007). *Inventing temperature: Measurement and scientific progress*. Oxford: Oxford University Press.

Daston, L., & Galison, P. (2007). *Objectivity*. New York: Zone Books.

Descartes (1979). *Discourse on Method.* (transl. D.A. Cress). Indianapolis: Hackett Publishing.

Descartes (1996). *Meditations on first philosophy.* (transl. J. Cottingham, rev. ed.) Cambridge: Cambridge University Press.

Forman, P. (2007). The primacy of science in modernity, of technology in postmodernity, and of ideology in the history of technology. *History and Technology, 23*, 1–152.

Friedman, M. (2010). A post-Kuhnian approach to the history and philosophy of science. *The Monist, 93*, 495–515.

Galison, P. (1997). *Image and logic*. Chicago: University of Chicago Press.

Galison, P. (2016). The pyramid and the ring: A physics indifferent to ontology. In B. Bensaude-Vincent, S. Loeve, A. Nordmann, & A. Schwarz (Eds.), *Research objects in their technological setting*. London: Routledge. forthcoming.

Glennan, S. (1992). *Mechanisms, models, and causation*. Ph.D. Dissertation. Chicago: University of Chicago.

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.

Harré, R. (2003). The materiality of instruments in a metaphysics of experiments. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 19–38). Pittsburgh: University of Pittsburgh Press.

Heidegger, M. (1967). *What is a thing?* (transl. W.B. Barton & V. Deutsch). Chicago: Henry Regnery Company.

Heidegger, M. (1977). The question concerning technology. In D. F. Krell (Ed.), *Basic writings*. New York: Harper & Row.

Heinemann, A.-S. (2013). Kalkül der Logik und Logische Maschine: George Boole und William Stanley Jevons. In R. Krömer & G. Nickel (Eds.), *Siegener Beiträge zur Geschichte und Philosophie der Mathematik* (Vol. 1, pp. 4–78). Siegen: Universitätsverlag Siegen.

Heßler, M. (2013). Die technisierte Lebenswelt: Perspektiven für die Technikgeschichte. *Zeitschrift Geschichte in Wissenschaft und Unterricht, 64*(5–6), 270–283.

Hughes, T. (2004). *Human-built world: How to think about technology and culture*. Chicago: University of Chicago Press.

Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method.* Oxford: Oxford University Press.

Keller, E. F. (2000). Models of and models for: Theory and practice in contemporary biology. *Philosophy of Science, 67*, 72–86.

Kornwachs, K. (2012). *Strukturen technologischen Wissens: Analytische Studien zu einer Wissenschaftstheorie der Technik*. Berlin: Edition Sigma.

Kroes, P., & Meijers, A. (2006). The dual nature of technical artefacts. *Studies in History and Philosophy of Science, 37*(1), 1–4.

Lenhard, J. (2011). *Mit allem rechnen*. Habilitationsschrift: Universität Bielefeld.

Lenhard, J., Küppers, G., & Shinn, T. (Eds.). (2007). *Simulation: Pragmatic constructions of reality* (Sociology of the Sciences Yearbook). Dordrecht: Springer.

Leonelli, S. (2015). What counts as scientific data? A relational framework. *Philosophy of Science, 82*, 1–12.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

Morrison, M. (1999). Models as autonomous agents. In M. Morgan & M. Morrison (Eds.), *Models as mediators* (pp. 38–65). Cambridge: Cambridge University Press.

Nordmann, A. (2006). Collapse of distance: Epistemic strategies of science and technoscience. *Danish Yearbook of Philosophy, 41*, 7–34.

Nordmann, A. (2012a). Im Blickwinkel der Technik: Neue Verhältnisse von Wissenschaftstheorie und Wissenschaftsgeschichte. *Berichte zur Wissenschaftsgeschichte, 35*(3), 200–216.

Nordmann, A. (2012b). Object lessons: Towards an epistemology of technoscience. *Scientia Studiae: Revista Latino-Americana de Filosofia e História da Ciência, 10*, 11–31.

Nordmann, A. (2013). Metachemistry. In J.-P. Llored (Ed.), *The philosophy of chemistry: Practices, methodologies, and concepts* (pp. 725–743). Newcastle: Cambridge Scholars.

Nordmann, A. (2015a). Review of “Léna Soler, Sjoerd Zwart, Michael Lynch, and Vincent Israel-Jost (eds.) *Science after the practice turn in the philosophy, history, and social studies of science*”, *Notre dame philosophical reviews: An electronic journal*, 21 June 2015, https://ndpr.nd.edu/news/58957-science-after-the-practice-turn-in-the-philosophy-history-and-social-studies-of-science/

Nordmann, A. (2015b). *Werkwissen oder how to express things in works* (Jahrbuch Technikphilosophie, Vol. 1, pp. 81–89). Zürich: Diaphanes.

O’Malley, M. (2011). Exploration, iterativity and kludging in synthetic biology. *Comptes Rendus Chimie, 14*, 406–412.

Roush, S. (2005). *Tracking truth: Knowledge, evidence, and science*. Oxford: Oxford University Press.

Schwarz, A., & Krohn, W. (2011). Experimenting with the concept of experiment: Probing the epochal break. In A. Nordmann, H. Radder, & G. Schiemann (Eds.), *Science transformed? Debating claims of an epochal break* (pp. 119–134). Pittsburgh: Pittsburgh University Press.

Tal, E. (2016). Making time: A study in the epistemology of measurement. *The British Journal for the Philosophy of Science*, *67*, 297–335.

van Fraassen, B. (2008). *Scientific representation: Paradoxes of perspective*. Oxford: Oxford University Press.

Waters, C. K. (2008). How practical know-how contextualizes theoretical knowledge: Exporting causal knowledge from laboratory to nature. *Philosophy of Science, 75*, 707–719.

Winsberg, E. (2010). *Science in the age of computer simulation*. Chicago: University of Chicago Press.

Wittgenstein, L. (1993). Appendix C: How can ‘knowing’ fit a physical fact? In *Philosophical Occasions: 1912–1951* (pp. 422–426). Indianapolis: Hackett Publishing.

Wittgenstein, L. (1994). *Philosophische Betrachtungen*. Wien: Springer.

Alfred Nordmann Dr.phil. in Philosophy from Hamburg University. Professor of Philosophy, Technische Universität Darmstadt. Research interests: Philosophy of technology as theory of scientific knowledge production, problems and methods of science and technology assessment, Heinrich Hertz and Ludwig Wittgenstein, working knowledge and theories of technologies art. Editor of book series *History and Philosophy of Technoscience*, co-editor of the *Jahrbuch Technikphilosophie*. Current edited volumes *Reasoning in Measurement* (with Nicola Mößner) and *Research Objects in their Technological Setting* (with Bernadette Bensaude-Vincent, Sacha Loeve, Astrid Schwarz).

Chapter 8
Constructive Philosophy of Technology and Responsible Innovation

Philip Brey

Abstract This essay argues for a new turn after the empirical turn in the philosophy of technology: the societal turn, which is the turn from reflective philosophy of technology (academic philosophy concerned with analysis and understanding) to constructive philosophy of technology (philosophy that is directly involved in solving practical problems in society). The essay aims to describe in detail what a constructive approach would look like and how it could be achieved. It claims that at least in the European Union, the conditions for a constructive philosophy of technology are favorable, due to the emergence in both policy and academics of the notion of Responsible Research and Innovation (RRI). It then goes on to describe how a constructive philosophy of technology can contribute to better technology development, better technology policy and better implementation and use of technology, through engineering-oriented, policy-oriented and use-oriented approaches to research.

Keywords Constructive philosophy of technology • Empirical turn • Societal turn • Responsible research and innovation

8.1 Introduction

In this essay, I will argue for a new turn after the empirical turn: from reflective to constructive philosophy of technology. I will relate this new approach to that of Responsible Research and Innovation (RRI), which has in recent years become important in both academic research and policy. My argument will be that it is possible to develop the philosophy of technology in such a way that it goes beyond academic analysis to become part of actual practices and policies of changing and

---

P. Brey (✉)
University of Twente, Enschede, The Netherlands
e-mail: p.a.e.brey@utwente.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_8

improving technology and its impact of society. This requires that the philosophy of technology is practiced in a way that is more oriented towards policy and engineering practice, and that involves collaborations and partnerships with other disciplines and with non-academic actors. The emerging consensus on the notion of responsible research and innovation in academics and policy offers an opportunity to give shape to a constructive approach to the philosophy of technology.

In the remainder of this introduction, I will explain the notion of constructive philosophy of technology and will defend its importance and feasibility. Having done that, I will conclude with a preview of the remainder of my argument in this paper. Let us begin with an explication of the notion of *constructive philosophy of technology*. This notion is defined in contrast with *reflective philosophy of technology*. I define reflective philosophy of technology as the academic study of technology and its relation to society and the human condition. This approach involves conceptualization, analysis, explanation and evaluation of technologies and engineering practices and their societal implications, and yields academic publications in which the results of these investigations are presented. The aim of this approach is to gain a better understanding, in general terms, of technologies, their impacts, and normative implications.

Constructive philosophy of technology does not have as its main objective to study technologies and gain a general understanding of them and their societal implications, but rather to help create better technologies with more beneficial implications for society. It sees as its task the development of philosophical ideas and approaches that come to guide and transform the practices of those actors in society that are responsible for the development, regulation and use of technology. It is an approach that is focused on specific problems and issues relating to technology and develops constructive and workable solutions for better development, implementation, use or regulation of technology. These solutions are often developed in direct collaboration with societal actors, both academic and non-academic, involved in the study, development, use and regulation of technology. Constructive philosophy of technology is thus activist and interventionist in a way that reflective philosophy of technology is not, and is directed at the development and application of concepts, principles and methods that result in knowledge that can be directly used by actors involved in the development, use and regulation of technology.

For many of those working in reflective philosophy of technology, changing and improving technology and its place in society is an ulterior aim of their studies. The academic publications that they produce may indeed contribute to such change in an indirect manner. However, they do not improve technology in a direct way, as they usually do not speak directly to those who are involved in developing or regulating technology and others who shape its impacts, and they do not usually include specific, directly usable knowledge or prescriptions that these actors can use directly in their own professional activities. Constructive philosophy of technology has therefore a direct role in shaping technology and its impacts that reflective philosophy of technology does not.

I do not use the label "applied philosophy of technology" to refer to approaches with a direct focus on changing and improving technology, because the whole notion of "applied philosophy" is problematic, as is the notion of "applied science". Philosophers of technology have long resisted a characterization of engineering as applied science, arguing that the engineering sciences centrally involves the development of new knowledge rather than the mere application of laws and concepts that were developed in the sciences (Skolimowsk 1966). In a similar fashion, I would like to maintain that constructive philosophy of technology centrally involves the development of new concepts and approaches, even though it will probably also make regular use of concepts and ideas that have been developed in reflective philosophy of technology.[1]

In arguing for a constructive philosophy of technology, I do not want to make the case that reflective philosophy of technology should be replaced by it. The reflective approach has its value as a general way of studying technology and its relation to society that provides broad insights and evaluations that have multiple uses in society. Notably, the theories and analyses of reflective studies can be of great benefit in constructive approaches. I would want to maintain that it is in the continued development and combination of constructive and reflective approaches that the philosophy of technology can make progress; constructive approaches often require analysis and evaluations of technologies and their implications for societies, and for this, they can depend on reflective studies, and reflective studies can improve their analyses and evaluations through consideration of the case analyses and practical instruments developed in constructive approaches.

How does my distinction between reflective and constructive philosophy of technology relate to the distinction between classical and empirical philosophy of technology that is made by proponents of the empirical turn (Kroes and Meijers 2000; cf. Brey 2010a)? I claim that both classical and empirical philosophy of technology, as they have been defined, are mainly reflective in nature. *Classical philosophy of technology*, the dominant approach in the field until the 1980s, is clearly by and large a reflective approach. Studies within this paradigm tend to look broadly at technology and its implications for humankind, often not focusing on specific technologies or technological practices, but on technology in general. They tend to have a deterministic conception of the evolution of technology and the impacts it generates, and tend to be either pessimistic or optimistic about its implications. It is an approach that does not generally include attention to empirical detail or collaboration with other, more empirical disciplines. Well-known studies in this tradition, by authors like Ellul, Heidegger, Kapp, Marcuse and Jonas, present

---

[1] It should also be clear that by arguing for a *constructive* philosophy of technology, I am not necessarily advocating an approach that is (socially) *constructivist*. I am arguing that the philosophy of technology should be more constructive, in the sense of being more focused on changing technology rather than just understanding it. This does not necessarily imply the (social) constructivist view that knowledge, technology and reality are the product of social meanings and processes, and that the physical world plays a small or nonexistent role in shaping and defining them.

interesting analyses and evaluations of technology, but do not directly shape any new practices of technological actors.

The empirical turn in the 1980s and 1990s yielded philosophical approaches that were more empirically informed and more multidisciplinary, and that had as their object of analysis more specific technologies, practices and issues in society. These approaches tended to see technological change and technological impacts not as deterministic but as contingent on all kinds of social actors and influences of society. These post-empirical turn approaches were sometimes labelled as *empirical*, *or empirically informed*, *philosophy of technology*. Although these approaches have a greater focus on actual practices than classical approaches, are more multidisciplinary and pay more attention to the social actors that develop, use and regulate technology, they have so far for the most part resulted in reflective philosophical research, with as end result academic publications, most of them with a philosopher as single author, with no direct intervention into technological practices.

Empirical philosophy of technology has, however, made possible the emergence of constructive philosophy of technology. It introduced an empirical, multidisciplinary orientation with a focus on practices and social actors, which constituted preconditions for a constructive philosophy of technology to succeed. In a constructive approach, further steps are taken to introduce practical aims for the research, to forge intense collaborations with both academic actors from other disciplines and non-academic actors to reach these goals, and to develop not just analyses and evaluations of technology, but also constructive tools for intervention. In this way, constructive philosophy of technology is an evolution of empirical philosophy of technology.

Constructive philosophy of technology is still in its infancy, but there have been dozens of studies in it already, especially in Europe, where the intellectual landscape and funding context for it have been favorable. My argument for pursuing this approach is threefold: (1) it is desirable to have such an approach if possible due to its potential benefits for society as well as the future success of the field of philosophy of technology, (2) there have already been past research initiatives that show that such an approach is possible, and (3) the conditions for more such initiatives are favorable, especially in Europe. Its continued development will involve trial and error, as any new approach would, but the potential benefits make worth these investments.

In the remainder of this essay, I will elaborate in more detail what a constructive philosophy of technology would look like and how it can be attained. In the next section, I will further describe the approach of constructive philosophy of technology, including a description of the unique intellectual contribution it can make to society. In the section that follows, I will argue that at least in the European Union, the conditions for constructive philosophy of technology and the collaborative projects it entails are good, due to the emergence in policy and academic circles of the concept of Responsible Research and Innovation. The next three sections will specify three types of research that a constructive philosophy of technology can engage in: engineering-, policy- and use-oriented, and will describe ways in which these contributions can be made. These sections are then followed by a conclusion.

8.2 The Societal Turn

The turn from reflective to constructive philosophy of technology may be called the *societal turn*, as it aims for collaboration with, and influence on, a variety of social actors who are involved in technological practices. Making this societal turn includes several challenges for philosophers of technology, among them the challenge of negotiating and maintaining their disciplinary identity in the multidisciplinary activities that are involved, maintaining high standards for research, and navigating the extent to which their activities qualify as research, and engaging in activities other than research, such as implementation projects, trainings and public outreach.

For the philosophy of technology to have a positive influence on technology and its place in society, there are three classes of actors that it has to successfully address and influence. First, it has to address *technology developers*: the engineers, entrepreneurs, manufacturers, marketers and others who are responsible for technological innovation and the development of new technological products. To influence technology developers is to influence what technologies are developed and how they are designed. Second, it has to address *regulators*: governmental agencies, legislators, policy consultants and other policy makers who have a role in setting and enforcing policies for the development and use of technology. To influence regulators is to influence the policies that help determine how technologies are designed and used. Ideally, the philosophy of technology should not only address regulators themselves, but also those actors who exert influence over the regulatory process, such as interest groups, NGOs, the media, and the general public. Third, it has to address *technology users*: organizations and individuals that implement and use new technologies. This constituency is important to address because the impacts of technology are determined in large part by the way in which these technologies are used.

Constructive philosophy of technology aims to succeed in directly affecting the beliefs and practices of these three classes of actors. It does so by engaging in research activities and interventions of three corresponding types: engineering-oriented, policy-oriented and use-oriented. *Engineering-oriented research* is research that addresses and aims to shape the practices, methods, beliefs and goals of technology developers. It addresses technology developers mainly through research collaboration, publications in media for technology developers, and talks and trainings for technology developers. *Policy-oriented research* is research that aims to shape technology policy and the practices of policy makers. It addresses policy makers and policy-relevant actors (including the general public) through collaboration, publications, talks and trainings. *Use-oriented research*, finally, aims to exert influence over the way in which technological products are implemented and used in organizations and by consumers. It addresses technology users through projects, publications, trainings, and contributions to public discussions about the use of technology.

Before proceeding to specific ways in which the philosophy of technology may make the societal turn for each of the three classes of actors, let us first consider in general terms how the philosophy of technology may make a contribution to practical issues and problems in society. Why should we believe that the philosophy of technology is equipped to make such a practical contribution? I will argue that the philosophy of technology is in a position to contribute unique knowledge and skills that can help make major improvements to technology development, technology policy, and technology use and implementation.

The philosophy of technology can contribute to diminishing the lack of understanding of the relation between technology and society that currently prevails. This lack of understanding is partially the result of overspecialization. Engineers have a sound understanding of technology, but have usually received little training that gives them insight into social processes and human behaviour, let alone the relation between technology and such social aspects. Similarly, social scientists have been trained to study social and behavioural phenomena but usually have little understanding of technology. Policy makers often have a lack of understanding of technology and its relation to society as well.

This rift between the engineering sciences on the one hand and the social and behavioural sciences and policymaking on the other makes it difficult to successfully develop, regulate and utilize technology in a way that takes into account its societal impacts and steers them into a more desirable direction. It is well possible that this rift constitutes a major reason why large-scale technological innovation projects fail, why the social consequences of technology are misjudged, and why opportunities in solving social problems are missed because those in charge do not know what the technological possibilities actually are.

Overcoming this rift requires interdisciplinary or transdisciplinary knowledge that transcends or synthesizes the vocabulary of the engineering sciences and the social sciences, and effective models for successful multidisciplinary collaboration between natural and technical scientists and social scientists. What is dearly needed is knowledge between the engineering sciences and the social sciences that will enable us to discuss the relationship between technology and society, technology and culture, technology and norms and values, technology and human behaviour, and technology and social needs. This knowledge can help give direction to the development and application of technology in society. Although it is not the only field that generates such knowledge, philosophy of technology has developed such knowledge over a broad spectrum and can in this way help bridge the gap between social science and engineering.

The research methods available to the philosophy of technology that give it a unique position in the analysis of technology and society include philosophical analysis, synthesis, and normative research. Its method of *synthesis* is the combination of conceptual frameworks, theories, paradigms and worldviews into larger systems by which hitherto disjointed phenomena can be understood relative to each other and as part of greater, meaningful wholes. Like no other discipline, philosophy investigates the relationship between fundamental and often abstract issues that cannot be easily investigated using empirical means, such as the relationship between language and reality and between science and religion. The method of synthesis enables investigations into technology that include a broad view and a broad agenda and identify the more abstract relations between technology and social phenomena. In this way, philosophy of technology can help our understanding of how technology relates to society, how the engineering sciences relate to the social sciences, and how these relations can be improved.

Philosophical *analysis*, a second method, is directed at attaining a better understanding of phenomena by conceptualizing them in a very clear and precise manner and by analysing their parts and the relations between them. Philosophical analysis is based on the idea that people's concepts, beliefs and reasons that they use to understand reality are frequently vague, incoherent or unsupported by reasons and can be improved through careful scrutiny and analysis. Applied to technology, philosophical analysis can help clarify the precise meaning of key concepts like "technology", "technological artefact", "sustainable development", "privacy" and "social impact", and can help understand and evaluate beliefs, theories, arguments and debates in the engineering sciences, social sciences, and in policy and public debate.

Third, *normative research* methods in philosophy consider how the world *should* be and how people *should* conduct themselves. Normative research is special in that it does not describe or explain reality, as most forms of research do, but prescribes how it should be. It does so on the basis of values and norms that define what is good and what we should strive for. Normative research takes place in ethics, which investigates how we should conduct ourselves and what are the conditions of a good life, but also in epistemology, which seeks to identify norms and standards for knowledge; in aesthetics, which investigates conditions for beauty and art; in political philosophy, which investigates how states and societies should be organized and how they should act; and in axiology or theory of value, which investigates which values should be most important to us. Normative research methods can be useful in solving social problems that involve technology by investigating values that are involved and ways in which these values are promoted or harmed. It can then assess and evaluate solutions, including technological solutions, relative to their expected consequences for the realisation of the values that it has found to be important.

The methods of philosophical synthesis, analysis and normative research enable the philosophy of technology to make a unique contribution not only to the study of technology and its relation to society, but also to improving the way in which technology is developed, used and regulated. However, for philosophy of technology to go beyond mere studies of technology and to be involved in this constructive role, there must be conditions present for it to do so successfully. These are not just conditions internal to the field itself, but also conditions in society. Specifically, there should be institutional structures in place that enable and support collaboration and interaction between philosophers of technology and the social actors that engage in technological practices. Such structures may include policies, organizational structures, and funding streams, amongst others. In the next section, I will argue that the efforts in the European Union to develop a framework for Responsible Research and Innovation currently offer the societal conditions for philosophers of technology to play this role. The European approach could possibly serve as a model for other parts of the world.

8.3 Responsible Research and Innovation

Responsible Research and Innovation (RRI) is an approach to research and innovation that has in recent years become an important component of European Union (EU) research and innovation policy (Owen et al. 2012, 2013; Van den Hoven et al. 2014). The term has become prominent in EU discourse since around 2010. It is the incarnation of a longstanding goal in EU policies to stimulate greater responsiveness of science and innovation towards society's needs. Research and innovation policy is seen in the EU as a means to promote its social and economic agenda. There has been a conviction in the EU that too much research is driven by intellectual curiosity only and not by the needs of society, and that too much innovation is driven by profit motives and does not respond to real needs in society. Since it is desirable that research and innovation help meet society's social and economic needs, and since it is believed this cannot be left to universities and the market, the EU has put strategies in place at the policy level that help orient research and innovation processes towards societal needs as defined in its social and economic policies.

One of these steps by the EU is to make its research and innovation agenda part of its social and economic agenda. In the EU's economic agenda, the so-called Europe 2020 strategy for the period 2010–2020 that aims at "smart, sustainable and inclusive growth" (European Commission 2010), research and innovation activities are defined as important means towards securing growth, and in ensuring that such growth is sustainable and takes into account social goals such as social integration and poverty eradication as well. The social agenda of the EU is incorporated into research and innovation policy amongst others by orienting a large part of the billions of EU research funding in its Horizon 2020 funding program towards societal challenges relating to health, ageing, well-being, security, sustainability, and social inclusion.

The RRI framework in EU research and innovation policy is in some ways the culmination of these initiatives. RRI is defined by the European Commission, the executive branch of the EU, as follows:

> Responsible Research and Innovation means that societal actors work together during the whole research and innovation process in order to better align both the process and its outcomes, with the values, needs and expectations of European society. (European Commission 2012, p. 2)

Philosopher and EC policy officer René von Schomberg has provided a frequently cited definition of RRI that goes into a bit more detail:

> Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical)

> acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society). (Von Schomberg 2012, p. 9)

Two aspects of RRI should be distinguished: its goal, which is, in short, to take into account societal needs as well as ethical criteria in research and innovation, and the way in which this goal should be reached. It is particularly about the latter aspect that many views exist. However, there is a fair amount of consensus that RRI can only be realized if societal actors work together more and are stimulated to take societal needs and ethical criteria into account, to the extent that they do not already do so. It is also held that these societal actors should include all relevant stakeholders, most centrally government actors, industry, universities, civil society actors (non-governmental organisations that represent the interests of citizens), as well as the general public.

RRI is currently a key component of EU research and innovation policy and is especially manifested in its research funding program, Horizon 2020, in which it is a cross-cutting theme. This means that the EU strives to incorporate RRI in many if not all research projects that it funds. In practice, this means that part of the funding in a research proposal is reserved for consideration of ethical aspects and/or engaging social actors with the project and/or other activities aimed at better aligning the research with societal needs. In addition to RRI being a key component of EU policy, several European states also have policies in place to support RRI, or some version of it, in their national research and innovation policies.

RRI offers philosophers of technology opportunities to be involved in multidisciplinary research and innovation projects in which they can be involved in helping to solve social problems and in making technology more responsive to societal needs and better in line with ethical criteria. The kinds of projects and activities supported in EU research policy enable both engineering-oriented contributions, for projects that are targeted at new technologies and technological innovations, policy-oriented contributions, since the EU also funds projects and activities that help it in devising better policies, and use-oriented contributions, since projects it funds include ones that either address the use of technology in society or in organizations or depend for their success on the successful introduction of a technology into society. This is not just a theoretical observation: many dozens of such projects have already been initiated or concluded in Europe in which philosophers of technology have played a role.[2]

---

[2] It should be cautioned, however, that no studies have been done of the effects of having philosophers in these programs and the degree to which they helped improve the outcome of them.

8.4 Engineering-Oriented Philosophical Research

Having observed that, at least in the European Union, the conditions are in place for the emergence of a constructive philosophy of technology, I will now turn to the three classes of constructive philosophy of technology that I distinguished earlier, and will consider in more detail how they may be approached. The first I will consider is engineering-oriented research, which was defined as philosophical research that addresses and aims to shape the practices, methods, beliefs and goals of technology developers. Engineering-oriented philosophical research can take various directions: it can be research in the philosophy of engineering that aims to conceptualize good engineering science and good design practice; it can be investigations into the nature of technological artifacts and the relation between design features and implications for users and for society; it can be investigations into the ethical specifications that designed products should meet and how these can be met in engineering design methodology; or it can be investigations into the professional responsibility of engineers and other technology developers. I will now highlight three specific approaches that are particularly promising, all in the realm of ethics.

One promising approach is that of *value-sensitive design* (VSD) (Friedeman et al. 2006; Van den Hoven and Manders-Huits 2009; Brey 2010b), an approach for designing technological products and systems in such a way that they conform to a desired set of (moral) values. Elaborate VSD methodologies have been developed to integrate considerations of value into the design process. The underlying assumption of VSD is that designed artifacts are not morally neutral but harbor tendencies to promote certain values or norms, or to violate them. For example, web browsers and apps may be either designed to protect the user's privacy, or they may offer no such protection or even actively violate user privacy. As another example, an ATM may be designed to be usable by all users, including the blind and people who speak different languages, or they may only be usable for those who have the required linguistic, bodily and sensory abilities, which goes against the value of universal access and, perhaps, of fairness.

VSD is a design methodology that involves identification of relevant values, translating them into design requirements and design features, and doing so in a way that is sensitive to contexts of use and that makes appropriate trade-offs between values. In this way, VSD is one of the first approaches that takes values seriously in design processes and that presents methods for systematically taking them into account. Many engineers, however, are still unfamiliar with VSD even the very idea that values can be included in designs. VSD offers a great possibility for ethicists of technology to collaborate with engineers to incorporate VSD methodology into specific design and innovation projects as well as to collaborate on incorporating VSD into standard design methodology in various engineering fields.

Ethical impact assessment (EIA; Wright 2014, 2011) has a broader scope than VSD. It is directed not only at technological design processes but also at larger innovation and infrastructural projects. EIA is not so much a methodology for incorporating values into design processes as it is one for assessing the ethical issues that may result from the project as currently planned. However, EIA is also concerned with taking mitigating measures if serious ethical issues or risks are identified. Wright defines an EIA as "a process during which an organization, together with stakeholders, considers the ethical issues or impacts posed by a new project, technology, service, program, legislation, or other initiative, to identify risks and solutions" (Wright 2014, p. 163). Unlike VSD, EIA centrally includes stakeholders in the process, and has its focus on the assessment of (proposed) designs rather than on the design process itself.

An EIA is similar to a social impact assessment or environmental impact assessment. It is a way of determining and assessing a project's implications for society. EIA focuses on ethical implications, such as whether a new information service that is being developed will sufficiently safeguard privacy and freedom of expression, or whether a new building will be designed in such a way that it upholds values like public accessibility, sustainability, and safety and security. An EIA's main steps include preparing an EIA plan, identifying stakeholders, consulting with stakeholders and analyzing ethical impacts, identifying risks and possible solutions, formulating recommendations for actors involved in the project that help mitigate unethical impacts and risks thereof, and working with these actors to implement these recommendations. Organizing and managing an EIA is another way in which philosophers of technology can collaborate with technology developers in projects.

A third approach is that of *ethical parallel research* (Van der Burg and Swierstra 2013), also called *embedded ethics*. In this approach, an ethicist becomes part of a technological project, interacts with engineers to learn from their research, performs ethical analyses of the research and the new technology that is being developed, and helps the researchers identify and deal with ethical issues in their research. Ethical parallel research could take the form of an EIA or employ VSD methodology, but it need not do so, and can use any kind of approach to address ethical issues.

These are three major ways in which the philosophy of technology can be directly involved in technology development. All three involve collaborations with technology developers in projects. This requires a mutual willingness of philosophers and technology developers to engage in such collaboration. As said, the RRI framework in the EU supports such collaborations. From philosophers, such collaborations may require new knowledge and skills, including a more than superficial understanding of how technology development projects work, what the technologies are that are being developed, and how their own contribution can be useful for technology developers. But if they can overcome these challenges and can succeed in assisting technology developers to recognize and address issues of ethics and valuation in their work, it is likely that products and services that are developed will make a better fit with morality and the values and needs of society.

8.5 Policy-Oriented Philosophical Research

Public policy consists of laws, mandates, regulations, funding priorities and courses of action initiated by governments to further their objectives, which normally include maintaining order and security, stimulating economic development, promoting the general welfare, and establishing justice. Although public policy is developed by government agencies and legislative bodies, non-governmental organizations, citizen groups and companies may also lobby for particular policies. There are two kinds of public policy to which philosophy of technology can contribute: technology and innovation policy, and policies for specific social and economic domains in which technology is involved. The latter category includes amongst others economic policy, environmental policy, health policy, educational policy, and social policy. The philosophy of technology can contribute to both types of policy by helping to ensure that policies incorporate a theoretically and empirically adequate conception of technology and its dynamics and impacts, and by helping to incorporate normative political and ethical analyses of technology. Such contributions will involve collaborations with policy makers and scholars working in governance studies, particularly the area of technology governance (Edler et al. 2003).

An important way in which philosophers of technology can aid public policies that involve new technologies is through the *ethical assessments of new and emerging technologies*. When new technologies are in development, such as nanotechnology or synthetic biology, assessments are needed of ethical issues associated with them, and recommendations are needed of how to incorporate such assessment into policies. This is where philosophers can make a contribution. Such assessments identify potential ethical issues with new technologies and with applications that may result from them, and may also suggest ways of mitigating or avoiding such issues through policies. The approach of *anticipatory technology ethics* (ATE) that I have developed (Brey 2012) is suited for such analysis at the policy level, although it can also be used for engineering-oriented analysis in a way similar to EIA. ATE can be used to make broad ethical assessments of new technologies and the artifacts, uses and impacts that may result from them. It employs futures studies and technology assessment to make projections of possible and likely future developments and does broad ethical assessments to identify a large range of ethical issues at different levels of analysis.

A second way in which philosophers of technology can contribute to policy is by proposing *distributions of responsibility* for risks and harms resulting from the development and use of technologies (Doorn 2010; Van de Poel et al. 2015). It is part of the expertise of ethicists to analyze various notions of responsibility and to make and justify attributions of (moral) responsibility. In determining the effects of a technology on society, there are usually many actors involved, and so there is a need to distribute responsibilities over these actors for potential negative outcomes. For example, if a self-driving car causes an accident, what is the responsibility of manufacturers, engineers, owners, and others for it? Such responsibility assignments can be the basis for laws and other policies that assign liability.

Philosophers can also contribute by proposing *models for stakeholder involvement* in technology development or political decision-making about new technologies. Many philosophers of technology have argued for the democratization of technology (Winner 1995; Feenberg 1992), meaning that all who have a stake in the development of technology can exert influence over it. But democratization of technology requires realistic models for stakeholder involvement that uphold democratic values, and philosophers may be in a position to develop such models.

Philosophers of technology can also help explicate *the role of technology in policy*. Technologies can help realize policy goals but they can also thwart their realization. An understanding is therefore needed of these processes. Technologies can help achieve policy goals because they have political properties (Winner 1980) and because they are able to influence and steer people's attitudes and behavior (Latour 1992; Illies and Meijers 2014). For example, recent studies have considered so-called persuasive technologies that change the behavior of users through persuasion and social influence can be used to promote sustainable consumptive practices (Verbeek and Slob 2006). Policy makers can make use of these properties of technology for policy. At the same time, policy makers should be aware of subversive effects of technologies for their policy goals, and take actions to mitigate such effects.

8.6 Use-Oriented Philosophical Research

Technology users come in two basic kinds: *individual users* and *organizational users*. An organizational user is an organization, considered as an agent, which has adopted a particular technology.[3] Although the organization can be considered a user of the technology, there are also end-users of the technology, which are employees (and sometimes also customers) that make individual use of the technology. For example, when a hospital adopts a hospital information system, the hospital as organization is a user of that technology, but the end-users are the administrators and doctors working in that hospital. When an organization develops or acquires technology with the purpose of making it available to customers or clients for use, like an Internet provider or car rental business, the organization is not itself a user but is rather the proprietor or owner of the technology, and its customers are end-users.

The philosophy of technology can contribute to the successful *implementation of technology in organizations*, understood in terms of it being able to perform its intended function without disruptive side-effects. It can make this contribution due to its ability to theorize and analyze the impacts of technology, for example on work, employee's well-being, and organizational culture, and to consider ethical issues that are in play, such as autonomy, privacy, and fairness (Brey 1999). It can contribute to good choices in the initial adoption of a technology, as well as to good policies for its use in the organization. In the same way, the philosophy of technology

---

[3] In addition, both individual users and organizations can be organized into user groups.

can contribute to the successful *provision of technology by technology providers*. For example, it can consider social, cultural and ethical aspects of alternative technological arrangements made by internet service providers and help devise good arrangements as well as effective and ethical internet service agreements between providers and users (Vedder 2001).

Another contribution the philosophy of technology can make is to an understanding and evaluation of individual uses of technology, regarding societal implications, implications for the individual, and ethical aspects. For example, it can provide analyses of implications of the Internet for friendship and social relations, and of the moral aspects of using stimulants in sports. Such analyses cannot only help users make informed choices about their use of technology, they can also help organizations that are faced with private technology users on their premises (e.g., restaurants, schools, airports) to develop effective and ethical policies regarding these users and their technologies, and they can help technology developers and makers of public policy. In addition, technological artifacts and services that are widely available to the public, such as smartphones, cars, commercial drones, Viagra pills, and (in some countries) guns, are good topics for public debate initiated by or participated in by philosophers, because of the immediate effect they have on the way of life.

8.7 Conclusion: Further Steps

In this essay, I have argued for a new turn after the empirical turn in the philosophy of technology: the societal turn, which is the turn from academic philosophy to philosophy as a constructive enterprise that is directly involved in solving practical problems in society. It is what I have called the turn from reflective philosophy of technology to constructive philosophy of technology. The essay spells out in some detail what this approach would look like and how it is practically feasible. I have argued that at least in the European Union, the conditions for a constructive philosophy of technology are favorable, due to the emergence in both policy and academics of the notion of Responsible Research and Innovation (RRI). Finally, I have described how the philosophy of technology can contribute to better technology development, better technology policy and better implementation and use of technology, through engineering-oriented, policy-oriented and use-oriented research, respectively.

Several challenges still lie ahead. Most importantly, as I have argued, a constructive philosophy of technology can only thrive in societies in which the appropriate conditions are present, which include policies, funding streams and alignments of actors that support the kind of multidisciplinary, applied, multi-actor research that is required for a constructive approach. Such conditions currently exist in the EU, but there are many countries where they are not present.

Even if these conditions are favorable, there are several other challenges. One challenge lies in ensuring that philosophers of technology who adopt a constructive approach have the required knowledge and skills to do so successfully. As claimed before, philosophers of technology will have to learn new knowledge and skills, including multidisciplinary skills, knowledge of nonacademic professional domains, and new philosophical approaches and methods. The many dozens of existing research projects that incorporate a constructive philosophical approach provide models of how to do this (or sometimes of how not to do it). In addition, though, philosophers of technology will have to experiment and develop their own techniques and approaches. The constructive approach does not come pre-packaged but will have to be developed in a process of trial and error.

Another challenge is for the field of philosophy of technology to find an adequate balance between reflective and constructive research. In a society in which most of the philosophical research is reflective, the opportunity is missed for the field to make direct contributions to society. But conversely, if most of the philosophical research is constructive, it risks becoming intellectually impoverished, because it cannot sufficiently feed on reflective studies. There is actually a risk in several EU countries for such a situation, since much of the funding for fundamental, reflective research has dried up in them, and most of the funding that is available is for multidisciplinary, applied projects.

A final challenge is that of maintaining critical distance. If philosophical research is undertaken in collaboration with often powerful societal actors, and is even funded or co-funded by such actors, including for-profit companies or government agencies that aim to further certain policy objectives, there is the risk that the philosophical research will adapt itself to the goals and views of these actors. It will be difficult to strongly argue against central practices in genetic modification if one's research is (co-)funded by genetic modification firms and involves collaboration with genetic engineers, or to severely criticize government policies on climate change if one's work is part of a government-funded multidisciplinary project to address climate change. Better safeguards are needed to protect the independence of philosophers and other scholars who participate in such projects, and there should always be enough funding available for truly independent, critical research.

Even if all these challenges are overcome, there is not yet a guarantee that a constructive philosophy of technology will actually be effective in addressing social and ethical problems that other fields have not been able to adequately address in the past. Yet, there are hopes that the philosophy of technology can do so, because it is special as a field in that it adopts a broad agenda regarding technology and its role in society, unique methods of philosophical analysis and synthesis, and a unique emphasis on (ethical) normativity. A turn towards constructive philosophy of technology is already occurring in many countries. I hope that this essay will help create awareness of this turn and will inspire more dialogue in the philosophy of technology on the benefits and pitfalls of constructive approaches and how to develop them in the best way possible.

References

Brey, P. (1999). Worker autonomy and the drama of digital networks in organizations. *Journal of Business Ethics, 22*, 15–25.

Brey, P. (2010a). Philosophy of technology after the empirical turn. *Techné: Research in Philosophy and Technology, 14*(1), 36–48.

Brey, P. (2010b). Values in technology and disclosive computer ethics. In L. Floridi (Ed.), *The Cambridge handbook of information and computer ethics* (pp. 41–58). Cambridge: Cambridge University Press.

Brey, P. (2012). Anticipatory ethics for emerging technologies. *Nanoethics, 6*(1), 1–13.

Doorn, N. (2010). A procedural approach to distributing responsibilities in R&D networks. *Poiesis & Praxis. International Journal of Technology Assessment and Ethics of Science, 7*(3), 169–199.

Edler, J., Kuhlmann, S., & Behrens, M. (Eds.). (2003). *Changing governance of research and technology policy: The European research area*. Cheltenham: Edward Elgar Publishing Limited.

European Commission. (2010). *Europe 2020: A strategy for smart, sustainable and inclusive growth*, Brussels, COM(2010) 2020 final. Online at http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf

European Commission. (2012). *Responsible research and innovation.* European Commission publications office. Online at http://ec.europa.eu/research/science-society/document_library/pdf_06/responsible-research-and-innovation-leaflet_en.pdf

Feenberg, A. (1992). Subversive rationalization: Technology, power, and democracy. *Inquiry, 35*(3–4), 301–322.

Friedman, B., Kahn, P., & Borning, A. (2006). Value sensitive design and information systems. In P. Zhang & D. Galletta (Eds.), *Human-computer interaction in management information systems: Foundations* (pp. 348–372). Armonk: M.E. Sharpe.

Illies, C. F. R., & Meijers, A. (2014). Artefacts, agency and action schemes. In P. Kroes & P.-P. Verbeek (Eds.), *The moral status of technical artefacts* (pp. 159–184). Dordrecht: Springer.

Kroes, P., & Meijers, A. (2000). Introduction: A discipline in search of its identity. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (pp. xvii–xxxv). Amsterdam: JAI.

Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. In W. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 225–258). Cambridge, MA: MIT Press.

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with ociety. *Science and Public Policy, 39*, 751–576.

Owen, R., Bessant, J., & Heintz, M. (Eds.). (2013). *Responsible innovation: Managing the responsible emergence of science and innovation in society*. Chichester: Wiley.

Skolimowski, H. (1966). The structure of thinking in technology. *Technology and Culture, 7*, 371–383.

Van de Poel, I., Royakkers, L., & Zwart, S. (2015). *Moral responsibility and the problem of many hands*. New York: Routledge.

Van den Hoven, J., & Manders-Huits, N. L. J. L. (2009). Value-sensitive design. In *Blackwell companion to the philosophy of technology* (pp. 477–480). Oxford: Blackwell.

Van den Hoven, J., Doorn, N., Swierstra, T., Koops, B.-J., & Romijn, H. (Eds.). (2014). *Responsible innovation 1: Innovative solutions for global issues*. Dordrecht: Springer.

Van der Burg, S., & Swierstra, T. (Eds.). (2013). *Ethics on the laboratory floor*. New York: Palgrave Macmillan.

Vedder, A. (2001). Accountability of internet access and service providers – strict liability entering ethics? *Ethics and Information Technology, 3*(1), 67–74.

Verbeek, P.-P., & Slob, A. (Eds.). (2006). *User behavior and technology development – Shaping sustainable relations between consumers and technologies*. Dordrecht: Springer.

Von Schomberg, R. (2012). Prospects for technology assessment in a framework of responsible research and innovation. In M. Dusseldorp & R. Beecroft (Eds.), *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden* (pp. 39–61). Wiesbaden: Vs Verlag.

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121–136.

Winner, L. (1995). Citizens virtues in a technological order. In A. Feenberg & A. Hannay (Eds.), *Technology and the politics of knowledge* (pp. 65–84). Bloomington: Indiana University Press.

Wright, D. (2011). A framework for the ethical impact assessment of information technology. *Ethics and Information Technology, 13*(3), 199–226.

Wright, D. (2014). Ethical impact assessment. In J. Britt Holbrook & C. Mitcham (Eds.), *Ethics, science, technology and engineering: A global resource* (2nd ed., pp. 163–167). Farmington Hills: Macmillan Reference.

Philip Brey Ph.D. in Philosophy at the University of California, San Diego. Currently full professor at the Philosophy Department of the University of Twente, the Netherlands and scientific director of the 3TU. Centre for Ethics and Technology. He is also current president of the International Society for Ethics and Information Technology, past president of the Society for Philosophy and Technology, and member of the editorial board of ten journals and book series in his field. His research focuses on Responsible Research and Innovation, ethics of emerging technologies, values in technology and design, and ethics of information technology.

Chapter 9
Towards a Third 'Practice Turn': An Inclusive and Empirically Informed Perspective on Risk

Rafaela Hillerbrand and Sabine Roeser

Abstract In this chapter we identify three practice turns in the social and philosophical study of technology that we also relate to risk analysis. The first practice turn singled out technology as a topic meriting serious investigation as a social phenomenon, the second turn steered the field towards the consideration of philosophical problems directly relating to what technology is and what engineers do. The third practice turn explicitly aims at changing the field's practice by close collaboration with the engineers. We argue that given the entanglement of evaluative and descriptive aspects of risk, it is important to develop approaches geared at this third turn, which is only now starting to take place. We propose that phronesis can play an important role in making context-sensitive assessments of evaluative aspects of risks, and that it can be assisted by emotions and art, as sources of moral reflection.

Keywords Risk • Ethics • Practice turn • Virtues • Emotions • Art • Value-sensitive design

9.1 Introduction: Engineering Ethics Today

> When philosophy paints its gray on gray, then has a form of life grown old, and with gray on gray it cannot be rejuvenated, but only known; the Owl of Minerva first takes flight with twilight closing in.
>
> G.W.F. Hegel, "Preface," *Philosophy of Right*

R. Hillerbrand (✉)
Karlsruhe Institute of Technology, Karlsruhe, Germany
e-mail: rafaela.hillerbrand@kit.edu

S. Roeser
Delft University of Technology, Delft, The Netherlands
e-mail: s.roeser@tudelft.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_9

In the famous dictum above, Hegel describes philosophy as reasoning that lags behind reality: only after the day is over, does the owl of Minerva, the symbol of philosophy, begin its flight. For a long time, ethics of technology seemed to be a paradigmatic example of Hegel's view of philosophy. Throughout most of the twentieth century, ethicists were content with a predominantly backward-looking evaluation of existing technologies. This also holds true for many approaches to technology assessment. There are some exceptions (e.g. Jonas 1979), but these are often so abstract that they fail to motivate policymakers or pragmatically-oriented engineers to reflect constructively on engineering design, or to embrace a more self-conscious form of technological progress guided by ethical considerations.

Today, at the beginning of the twenty-first century, ethics of technology and engineering have outgrown their traditional image as birds active only at twilight. Far from relying on retrospective assessment, contemporary analytically-oriented ethics of engineering aims to play a constitutive role in shaping technological progress. This new self-understanding is accompanied by changes in other areas of philosophy of technology, such as metaphysics and ontology. Contemporary philosophy of engineering sees technological design as an amendable process, where technology is a means to an end that usually arises outside the technology itself and hence opposes the technological determinism espoused by Heilbroner (1967) and others. Engineers often have a similar view of technological design. In engineering practice, economic and other external values (such as safety or efficiency) are often taken into account from the outset of the design process. The next step is to integrate moral and social values right from the start of the technological development.

Today's engineering ethics aims to integrate ethical and social values into early stages of product development, when many design-related and institutional aspects of the technical artifact or system are still malleable. This integrative method appears in various design-for-value approaches (van den Hoven et al. 2015) and in forward-facing technology assessment in which ethical aspects are central (Decker 2013). It requires interdisciplinary efforts and a combination of empirical and normative reasoning. Thus a prospective ethics of technology is part of broader trends within contemporary philosophy of science and technology, where the so-called 'practice turn' has underscored the need for dynamic exchange between philosophy and the specialized sciences (e.g. Soler et al. 2014).

The social and philosophical study of engineering and technology has actually seen several practice turns over the last decades.[1] The first practice turn singled out technology as a topic meriting serious investigation as a social phenomenon. This turn relied on the methods of the social scientists who initiated it, and it gave rise to the field of STS (Science and Technology Studies). The second practice turn began with Kroes and Meijers (2000) – more specifically, with what they called the 'empirical turn' in philosophy of technology. It steered the field away from broad abstract reflections on technology as a general phenomenon and toward the

---

[1] While these are often referred to as 'empirical turns' in technology studies, we prefer the term 'practice turn', which brings us into line with other reflective disciplines such as sociology or philosophy of science.

consideration of philosophical problems directly relating to what technology is and what engineers do.

In recent years, an aim to impact on the engineering design process itself and develop new quality standards for engineering processes has emerged within philosophy of engineering. We contend that this constitutes a third practice turn. An integral tactic of the second turn was to base philosophical reasoning on practices in the field. Today, scholars reflecting on technology and engineering increasingly hope not only to study the field but also to change its practices. This third practice turn builds on the first two turns and engages with the work of the designers and developers it aims to affect. The so-called value-sensitive design approach illustrates this very clearly: philosophical-ethical conceptual considerations are brought forward with the explicit goal of including ethical values in early stages of the design process (third turn). These considerations are fine-tuned by studying the designers' work in close collaboration with the designers themselves (second turn). This is combined with empirical studies of the interests of the various stakeholders and user groups associated with the technology (first turn).

Along with value-sensitive design, responsible innovation provides a paradigm case for the third practice turn. And other areas of applied philosophical research, which cannot be subsumed under one of these categories, also aim to improve concrete technological products or processes. For example, improved criteria for sustainability analysis are derived within a certain ethical framework (Reitinger et al. 2011, 2012; Maga 2015; Künneke et al. 2015). More generally, prospective technology assessment does not content itself with evaluating existing practices but attempts to change them according to certain ethical and social standards (Decker 2013).

But do all these approaches really indicate, as we suggest, something new – a third practice turn? In order to address this question, in the following we will zoom in on risk analysis as a central field of technology and engineering studies. We will argue that the third practice turn (as defined above) requires a crucial shift in contemporary risk analysis. This shift involves an even more integral treatment of both the normative and the descriptive aspects of risk right from the very start. In Sect. 9.2, we review the history of risk research over the last half-century or so. We highlight the impacts that the first two practice turns had on the field of risk studies and on our understanding of risk. Directions for a third practice turn are also identified. Section 9.3 argues that this third practice turn is necessary due to the inevitable entanglement of descriptive and normative elements in risk analysis. We contend that one cannot disentangle the ethical basis for making decisions about potentially risky technologies from the question of how to handle various uncertainties. Classical as well as most contemporary ethical approaches, however, treat uncertainty as a kind of secondary complication, an afterthought to moral judgments under certainty. In this chapter, we argue for a synthesizing approach that considers the interrelation of ethical aspects and uncertainty right from the start of a risk analysis. We maintain that this requires risk analysis to accomplish a third practice turn, much like the one outlined above for studies of technology and engineering more generally. Section 9.4 argues that one way to integrate analyses of the descriptive and normative aspects of risk is to revitalize the ancient concept of phronesis. Section 9.5 details how this virtue can be cultivated with the help of emotions and art.

9.2 Risk Analysis

Risk analysis originated during debates about the non-military use of nuclear power in the 1950s and 1960s. In these early days, risk analysis was mainly seen as a quantitative, mathematical and scientific endeavor, largely free from any normative or qualitative aspects. We refer to this as the technocratic approach to risk. Despite the potentially negative connotation of this terminology, we recognize and want to stress the great advances the technocratic approach made in the understanding of risk. This somewhat standard approach often defines risk as a product of probabilities and unwanted consequences. Unwanted consequences can be, for example, measured in terms of the number of deaths, as with annual fatalities resulting from a technology or activity; however, they can also include other effects on humans or nature. Risk analysts then try to single out those technological developments that minimize risk in order to determine which activities or technologies should be pursued. In cases where probability estimates are unreliable, various modifications to this simple approach have been suggested (e.g. Kaplan and Garrick 1981). But just like the standard approach, even these more sophisticated versions provide a limited perspective on risk. The field began to outgrow these restrictions in the late 1960s with the first empirical turn, when the social sciences – most prominently psychologists, sociologists, and economists – discovered (technological) risk as a topic for their own studies. The authors of the first turn critically engaged with the technocratic approach to risk. In 1969, Starr published a seminal work on acceptable risks that first introduced a less objective view of risk in which also subjective risk-preferences were considered (Starr 1969). Especially influential is the work of psychologists Kahneman and Tversky (e.g. Kahneman and Tversky 1974), for which Kahneman received the 2002 Nobel Prize for Economics. They investigate empirically how ‘real people’ perceive and deal with risks. They describe the ways in which real actors deviate from the ideal utility-maximizing rational individual (the technocratic ideal) when it comes to risky and uncertain decisions. The initial idea was that when people’s risk perceptions deviate from rational decision theory, they are mistaken. Kahneman (2011) and others think that this is because people often rely on heuristics that are prone to be biased (Gilovich et al. 2002).

The psychologist Paul Slovic has also studied the risk perceptions of laypeople empirically. Originally working along the same lines as Kahneman and Tversky, Slovic developed an alternative hypothesis: laypeople do not have a *wrong* perception of risk, but rather have a *different* perception of risk than experts do. When Slovic and his colleagues asked people to judge the *risks* of activities or technologies, laypeople’s estimates differed significantly from those of experts. However, when these laypeople were asked to judge *annual fatalities* due to activities or technologies, their judgments came close to those of experts. The technocratic definition of risk is based on the mean expected harm, i.e. possible harm multiplied by its occurrence probability. Here harm is parceled into straightforwardly measurable quantities, such as fatalities over a fixed time period. Experts tend to *define* risk as the expected number of fatalities in a certain time, whereas for laypeople, these are apparently separate notions. So laypeople attach different connotations to the notion of risk, which explains why they can perceive risks differently than experts do while still making estimates of annual fatalities that align with those of experts. Hence, for laypeople, risk seems to be much more than the product of harm and probability. Slovic and his colleagues conducted further studies in which they identified the additional considerations that are involved in laypeople's risk perceptions (Slovic 2000, p. 86). These include issues of fairness, equity, and control. More recently, Gerd Gigerenzer (2007) has also challenged the heuristics and biases approach by arguing that intuitive risk perceptions can be more reliable than the prescriptions of the technocratic approach (Gigerenzer and Gaissmaier 2011).

Due in part to a series of major technological disasters in the second half of the twentieth century, the 1980s witnessed rich scholarly debate about the social aspects of risk. Renowned sociologists have also contributed to a more qualitative approach to risk (e.g. Beck 1986; Giddens 1990; Foucault 1991). Raynor and Cantor (1987) focused on aspects of fairness in technological risk analysis, while Kasperson et al. (1988) assessed broader societal implications. Douglas and Wildavsky (1982) offered a particular cultural theory of risk. Since the 1990s, social aspects of risk have also been studied under the heading of governmentality, a term and concept borrowed from Foucault. Here institutions and organizations, and the different ways in which they organize power and govern populations, take center stage in risk discourse (Foucault 1991; O'Malley 1999). At roughly the same moment, Beck's and Giddens's idea of the "risk society" became popular amongst a broad range of scholars and laypeople alike (cf. Beck 1986; Giddens 1990). Beck's and Gidden's arguments differ, but they both use the term "risk society" to denote the phenomenon by which contemporary Western societies become more and more occupied with the uncertainty of the future. This leads to attempts to reduce this uncertainty, and to a fixation on hazards and insecurities. These are often studied with the help of risk theory.

These aforementioned psychological and sociological approaches constitute a first practice turn, as they focus on empirical reality rather than on the ideal-type reasoning of the technocratic approach. The second practice turn in risk analysis occurred in the late 1980s and early 1990s, when risk became a topic for philosophical reflection within ethics and decision theory, generating normative arguments that called for changes in standard approaches to risk. In the 1990s, pioneering work was done by Sven Ove Hanson and Kristin Shrader-Frechette (1991). Hansson discussed conceptual as well as ethical questions about how to deal with various types of uncertainties. Shrader-Frechette did seminal work in the context of risk and environmental ethics. In the new millennium, more philosophers have followed them by studying conceptual and ethical aspects of risk (Asveld and Roeser 2009; Roeser et al. 2012 provide overviews of the second practice turn). In these debates as well, what we referred to as the technocratic approach to risk figures prominently. Often, philosophers argue against it or call for it to be supplemented because it tends to overlook important ethical and societal issues, such as justice, fairness, autonomy, and responsibility. However, in the form of rational decision theory, the technocratic approach still features prominently even in today's philosophical analysis. Here the focus is often not so much on risk minimization as on maximizing expected utility. Expected utility is defined as the product of all (positive and negative) utilities and their occurrence probabilities. Expected-utility theorists frequently assign monetary values to the expected outcomes, and then compare costs and benefits on an aggregate level. This can be seen as an extension of the utilitarian paradigm to decisions under uncertainty. The expected-utility approach makes use of the same conceptual apparatus as the standard technocratic approach to risk, but unlike risk minimization (where risk is understood to mean harm) it does not emphasize the negative effects alone.

In the context of the second turn, some philosophers of risk focus on changing practices in the realm of policy-making (e.g. Janssen et al. 2005; Brook et al. forthcoming), but not the practices of engineers or scientists. However, we think that it is important to go a step further, by initiating a third practice turn in the study of risk, analogous to the third practice turn in the study of technology and engineering in general. While the second practice turn involves philosophers paying attention to the actual practices in these fields, critically reflecting on them, and providing arguments for changes in policy, the third practice turn requires philosophers to engage with and aim to change the practices of technology and engineering and of risk analysis as engineers or scientists have established it. The third practice turn that we call for in this chapter sees philosophical reasoning and risk analysis as fundamentally entangled. Instead of merely criticizing the technocratic approach to risk, philosophy of risk after the third practice turn should attempt to overcome such reductionist views of risk – views that do not pay explicit attention to a broader range of relevant evaluative issues – by engaging with science and engineering. It should provide for approaches to risk that do take ethical considerations explicitly into account.

Within the philosophical work on health care, this third turn is already visible. Some scholars aim to address ethical issues here in terms of risk; work on the risks of biomedical technologies by F. Steger provides an early example of the third practice turn (Steger and Hillerbrand 2013). In the philosophical study of risk in other than biomedical contexts, the third turn is less prominent, but some evidence of it appears in the general study of technological risks. Van der Poel's and Royakkers's textbook (2011), for example, teaches ethical reasoning to engineering students, and also addresses the topic of risk. This and similar work falls under the third turn as it aims to teach ethics to those who can incorporate ethical reasoning into early stages of product development. Other studies on risk directly engage with engineering work. For example, Taebi and Kloosterman seek not only to improve the discourse about the design of nuclear reactors and repositories of nuclear waste, but also to improve the design itself from an ethical point of view (e.g. Taebi and Kloosterman 2008).

9.3 Entanglement of Descriptive and Normative Aspects

In the following pages, we argue that a third practice turn is required in risk analysis because normative and descriptive aspects of risk cannot be fully disentangled. This requires experts to engage with ethically normative issues in the early stages of product development or conceptualization and we hold that this could be achieved by engaging philosophers. This in turn requires that philosophical analysis seeks an even closer exchange with the engineering practice. Our claim builds, firstly, on the recognition of risk as an inherently normative as well as a descriptive term. Secondly, scientific reasoning, at least in the applied sciences needed for risk analysis, is not free from ethical or social values. The second and the first turns, however, fail to fully address this entanglement.

9.3.1 Risk as a Thick Moral Concept

Though proponents sometimes praise the technocratic approach to risk analysis as an objective and value-neutral method (e.g. Sunstein 2005), the very act of deciding what counts as an ‘unwanted outcome’ involves an ethical judgment (Fischhoff et al. 1981; Jasanoff 1993; Slovic 1999). Hence, even the simplest concept of risk understood as mean harm has normative aspects. Technological risks can affect people’s wellbeing, which gives rise to ethical issues. The evaluative and descriptive aspects of risk cannot be disentangled, and in that sense risk is a so-called ‘thick concept’ (Möller 2012). This term originated from metaethics and refers to concepts that have empirical and ethical aspects at the same time.[2]

Lying and stealing are examples of thick concepts: they describe an action but also entail an ethical evaluation, and according to some scholars, these two aspects cannot be conceptually separated (cf. Little 2000; cf. the literature on the dual nature of artifacts in the context of technology; Kroes and Meijers 2006; Franssen 2006). Thick concepts differ from thin ethical concepts (like ‘just’) that are purely evaluative or normative. When judging a technology as ‘risky,’ speakers not only make a descriptive claim (for instance, the claim that an accident involving this technology may cause a large number of fatalities) but also express their evaluation (i.e. their moral judgment that this is a bad feature of the technology). Like risk, sustainability and safety are important thick concepts where technologies are concerned (cf. Hillerband 2015).

While the empirical aspects of risks can be investigated and measured by scientific methods, they are not sufficient to judge the ethical aspects of risks. Moral philosophers of various stripes agree that one cannot derive norms and values

---

[2] Cf. McDowell (1981) and Williams (2006). Though the distinction between thin normative concepts and thick ones may not be a sharp one, it is nonetheless conceptually useful (e.g. Kirchin 2013).

from facts, or ‘ought’ from ‘is’ (Moore 1988 [1903]). Empirical information about risks is necessary, but not sufficient, for an ethical assessment. Analyzing the ethical or evaluative aspects of risk requires ethical reflection. We need such reflection when considering the value of human life, the possible negative side effects of technologies (such as impacts on health and the environment), and how best to distribute risks and benefits. Even standard technocratic risk analysis makes significant ethical commitments. Yet the entanglement of risk’s ethical and descriptive aspects means not only that an ethical assessment is inevitable, but also that we cannot neatly separate descriptive and ethical assessments. For example, one might think that one could first conduct a technical and scientific assessment and then come up with an ethical evaluation. But this model does not acknowledge the evaluative commitments that are inherent in the method by which the risks are measured and compared in the supposedly purely descriptive assessment. By explicitly attending to their moral underpinnings, we can and must rethink conventional scientific and quantitative approaches to measuring and assessing risk. In the next section, we will discuss this in more detail.

9.3.2 Moral and Social Values in Descriptive Sciences

Identifying risk as a thick concept means that normative considerations, such as ethical and socio-political questions, are indispensable from the outset in risk analysis. This requires a radical paradigm shift in the current practice. In contemporary risk analysis, the sciences (including the social ones) assess the impact of technical interventions first; only in a second step are normative evaluations considered (cf. Hillerbrand 2011). Consider as an example Germany’s phasing-out of nuclear power. After the accident at Fukushima Daishi, the federal government set up the Ethics Committee for a Safe Energy Supply (Ethikkommission für eine sichere Energieversorgung) in order to assess the risks from Germany’s nuclear power plants. But instead of working hand-in-hand with the Reactor Safety Commission (Reaktorsicherheitskommission, RSK) that engages with the scientific and technological assessment of nuclear power plants, the ethics committee only played a secondary role in evaluating the RSK’s findings. Though its interdisciplinary personnel made the ethics committee a kind of ‘thick’ group, taking the “thick” nature of risk seriously would have required that ethical evaluation precede, accompany, and guide the RSK’s analysis. Such evaluation could have provided a list of issues and values to be considered in the descriptive analysis.

Beyond the insight that risk is a thick concept, recent discussions within philosophy of science highlight the value-laden nature of empirical research itself (e.g. Biddle and Winsberg 2010). This conversation dates back to earlier work by Rudner (1953) and Churchman (1948, 1956). While authors like Kuhn have long stressed the importance of epistemic values in the descriptive sciences, these scholars focus on moral and social values, which they argue cannot be completely eliminated from descriptive research. In the following pages, the term “values” refers to moral or social values. These ideas go far beyond Weber's original dictum that the scientist's motivation for a certain research project is driven by values. Rather, this research suggests that the scientist makes value judgments through applying the scientific method.

Research on the risks of new and existing technologies is a topic within applied science, and as such is a paradigm case of value-laden investigation. Let us therefore recap the arguments that philosophers of science give for the value-laden nature of empirical research. Roughly speaking, Richard Rudner and Charles West Churchman identified the aim of science as accepting (or rejecting) some hypothesis *H* based on inductive reasoning. They argued that there always remains a certain "inductive risk," because the question of how much empirical data is needed to accept or reject *H* cannot be answered with absolute certainty. The answer to this question depends, at least in part, on the potential implications of accepting a false hypothesis as correct (or a correct one as false). Just think of declaring a new vaccine to be safe and introducing it to the market, when in fact its side effects are more detrimental than the disease itself. The seriousness of such a misjudgment depends on normative assumptions, e.g. how one compares the risk of the potential deaths of those vaccinated to the possible benefits of curing the disease. In this way, value judgments sneak into the empirical sciences.

In 1956, Jeffrey famously turned against this type of argument. He started off with a rejection of these assumptions about how science works (Jeffrey 1954). For Jeffrey, science is not about refuting or accepting a hypothesis *H*; rather, its aim is to assign a genuine probability p (H|E) to a hypothesis *H* that is conditioned on empirical evidence E. It is then up to people other than the scientists to decide whether this probability is sufficient to support actions that are motivated by the assumption that *H* is correct. After Jeffrey, the power of probabilistic predictions has often been seen in the supposed disentanglement of value judgments on the one hand and empirical descriptive sciences on the other. This allows for a division of labor between the descriptive aspects of risk (done by scientists) and the normative evaluation of risk (done by policymakers, ethicists, and others). This division of labor still underlies contemporary policy-making when it comes to risky technologies.

But Jeffrey's argument was challenged by Heather Douglas (2000). We reconstruct her counterargument in the following way. Often, there exist various tests that will assign conditional probabilities to *H*. One of these tests may be more sensitive and will thus produce more false positives, while the other will produce more false negatives. The method you choose, the way in which you acquire and use empirical evidence *E*, depends on the impact that false negatives and false positives may have. Hence, the choice of the method requires a value judgment, as some of its effects are not ethically neutral. Jeffrey's argument for the value-neutrality of probabilistic forecasts thus does not hold, and the common division of labor between descriptive and normative reasoning breaks down.

But in order to address Douglas's concerns, can we modify Jeffrey's idea by including the empirical basis in the probabilistic information? This argument would build on the premise that science is not about assigning probabilities p(H|E) that are conditioned on the empirical evidence alone, but rather about assigning probabilities $p(H|E,E_1)$ that are conditioned on the empirical evidence and on the test E_1 used to extract this very information. We then, however, can again apply Douglas's argument to the probabilities conditioned on E and E_1: the way in which information from the available tests is used depends on how important false negatives and false positives are. Hence also Douglas can repeat her argument against Jeffrey as also here value judgments cannot be avoided.[3] Note that this does not imply that (applied) scientific analyses are flawed in any sense. It rather means that the descriptive sciences cannot avoid (mostly implicit) value-judgments.

9.4 Limits of Rule-Based Approaches

9.4.1 Entangling Ethical and Epistemic Aspects

In the preceding section, we provided two lines of reasoning showing why the ethical-normative aspects and the empirical-descriptive aspects cannot be completely disentangled in analyzing risks. But where to go from here? And how does this connect to the third practice turn that we want to call for?

Let us start by briefly pointing out how existing approaches to dealing with risk fail to account for its thick nature. Let us consider risk minimization as a variant of expected-utility approaches, as well as the precautionary approach. In a nutshell, proponents of both approaches seem to settle for the respective underlying ethical principle for decision under certainty, adding the "risky aspect" as a kind of secondary complication. The entanglement of normative and descriptive aspects (which essentially contain information about the uncertainty of the potential harm), however, renders this approach inadequate. Let us explicate this in a little more detail.

As detailed in the preceding sections, risk minimization as a variant of expected-utility maximization settles for a certain interpretation of a utilitarian ethics. Expected-utility analysis and common quantitative approaches to risk both involve some heavily contested normative and scientific presuppositions. For example, negative and positive implications are traded off against each other, though they may touch on incommensurable values. Incommensurability means that not even an ordinal ranking of the values is possible; no positive comparative judgment (e.g. no 'better than') of these values can be made (cf. Chang 1997). This can be partially accommodated in a multidimensional risk analysis. For example, the loss of human life or environmental damage is "traded off" against certain economic benefits. In fields such as bioethics, risk is therefore often defined using multiple categories. However, even these approaches judge a decision solely in terms of the harms and benefits that can be expected on average. When the people who bear the costs are

---

[3] Bayesian approaches can be seen as reasoning along these lines; however, they also face the problem that they cannot stop the regress by *formal* arguments alone (cf. Frisch 2015).

different from those who reap the rewards, the aggregated view seems problematic. Consider the case of a clinical trial. Under expected-utility theory, severe harm to the research subject may be acceptable when the mean expected benefit is high. However, considerations of fairness may well require that we put more weight on the research subject's utility than on what a person would suffer or gain "on average" – the average person being a mathematical construct. Similar problems arise when we consider the siting of large technological facilities, such as power plants. It is not the concern of this chapter to argue for or against utilitarian approaches. Instead, we want to point out that in risk or expected-utility analysis, a decision about a certain ethical framework has already been made before the reasoning about the genuinely risky aspects – the uncertainty – begins. One settles on an ethical evaluation (in this case a certain type of utilitarian analysis, i.e. cost-benefit analysis), and then uncertainties are added only as a secondary complication. This runs counter to the very nature of risk, in which uncertainty and ethical aspects cannot be disentangled.

Further problems with expected utility and risk analysis exist that are more epistemic in nature. Look at the following decision situations A and B under risk (cf. Roeser et al. 2015). If one decides to do A or B, in both cases the costs and benefits are borne by the same person; in both cases one knows the full probability distribution of the harm such a decision may cause. However, the probability functions differ between the cases. For A the probability function is Gaussian; for B it is a distribution with heavy weights in the tail, but with the same mean as the Gaussian distribution. Risk analysis focuses only on the mean – which implies that it would calculate the same expected utility in both cases. However, in the case of the fat tail distribution, rare events that may cause a great deal of harm are much more likely than in the case of the Gaussian distribution. Hence, it is not at all clear why one should use the mean as the only criterion on which to base one's decision. Risk minimization seems to be apt for decisions where all the information on the probability distribution is entailed by the mean, i.e. where the distribution is normal. But for many of the side effects of new technologies, there are no a priori reasons that this should be the case. This is particularly true when side effects touch on complex systems such as the biosphere or local economies. A decision procedure that incorporates not only the mean but also the higher moments of the distribution, thus providing information about the shape of the probability distribution, would be more adequate here. Note that technological threats are often of type B, as they result from complex interactions within a complex environment. When one is trying to get a stable power supply from a renewable resource such as wind, for example, sudden gusts constitute problems. The probability that these will occur is non-Gaussian, with heavy tails. As Böttcher Peinke et al. (2004) have shown, designing a wind turbine according to the standard approaches that look only at mean values assumes that wind gusts which actually occur on a weekly basis will only occur once a century. This example demonstrates that some risks cannot be captured adequately when only mean values are evaluated. More importantly, this example shows that the question of which decision criterion to use – one that takes into account the mean only or one that also accounts for higher moments of the probability distribution – cannot be disentangled from the ethical importance of the possible harm under consideration.

Now consider the precautionary approach as an alternative to dealing with risky technologies. Roughly speaking, the precautionary principle focuses on the worst-case outcome only and avoids it at all costs. This is in sharp contrast with risk minimization, which averages over all possible outcomes. If the worst-case scenario does not occur, then a precautionary strategy might be far from optimal. Consider an example from geoengineering, namely solar panels designed to counteract global warming by reflecting solar radiation. It may not be feasible to conduct extensive testing in order to completely rule out severe side effects, e.g. via global cooling, that could threaten the existence of human life on Earth. The precautionary principle (or minimax rule) would then forbid launching this new technology, no matter how high its possible benefits. It is also not an option to completely neglect the information about the uncertainty of potential harm, and to settle for elementary (i.e. non-probabilistic) decision criteria, as this is not adequate to the complexity of the situation. Disentangling ethical evaluation from this information about uncertainty is impossible.

While the second practice turn has taught philosophy to be aware of practices within engineering and science and to ascertain the involvement of normative judgments in the sciences, the third practice turn aims to feed the insights about value judgments gleaned from the second practice turn back into the practice itself in order to change that very practice. In other words, where the second practice turn makes philosophy of technology, science and risk more attentive to the empirical domain it investigates, the third practice turn seeks to make technology, science and risk analysis more attentive to their own inevitable normative foundations. We interpret the entanglement of the ethical and the descriptive aspects of risk as a need for a third practice turn. What this implies for risk analysis we adumbrate in the following subsection.

9.4.2 In the Need of Intellectual Virtues

Precautionary approaches, risk minimization and other common tactics for making decisions in the face of risks and uncertainties, all have their pros and cons when dealing with technological risks. There seems to be no decision approach that acts as a silver bullet for all the practical problems a prospective technology assessment faces. Which decision approaches prove most adequate depends on many context-dependent features. Formal approaches on their own inevitably fall short. Even though this fact is acknowledged by some scholars, formal approaches to risk are often used by practitioners and policymakers as supposedly objective methods that provide us with ultimate verdicts, irrespective of many context-dependent features.

However, a general solution to the problem of assessing risk, detached from contingent features of the decision situation, appears to be out of reach. As Luntley (2003, p. 326) notes:

> The ethically competent need general rules, but these are not what primarily lie behind ethical competence in decision making. Wise judgement is not constituted by grasp of general rules, but by the attentional skills for finding salience in the particularities of situations. The important element of decision making […] is the element that turns on the possession and operation of these attentional skills.

This type of judgment seems to be vital for all who make decisions about the design or market release of a new technology. Decision-theoretic approaches like those discussed in the last subsections, as well as more advanced versions provide helpful guidance. But more is needed, and we agree with Luntley that what is needed is a certain attentional skill. This skill can be further elucidated via the Aristotelian concept of phronesis (Hillerbrand 2010). Phronesis is a special dianoetic virtue, i.e. intellectual virtue where reason and argument are central. At the same time, phronesis is always attentive to the morally good. Since Antiquity, it has been argued that phronesis is concerned with particulars, as it provides guidance for action in concrete situations. One can learn the principles of action, such as which decision approaches can be applied to uncertain and risky technologies. But applying them in the real world requires experience of the world.

In doing so, we follow the ancient understanding of phronesis. It refers to a certain ability and willingness to (i) identify situations as morally relevant and to (ii) implement methods for realizing a moral norm in real-life situations. Thus a first task of phronesis is to identify certain decision situations as ethically relevant. This may be the market-launch of a new technical product or the very first design steps towards a new artifact. A second task is applying (general) rules. Thus phronesis arbitrates between general normative rules – as, for example, in precautionary theory or risk minimization – and a specific decision context. This second duty parallels to some extent the duty that, according to Kant, must be carried out by judgment, the ‘praktische Urteilskraft.’ Rehabilitating an ancient dianoetic virtue may help to solve problems of modern technologies. The central role of phronesis in an ethics of science, engineering and risk has already been emphasized by several authors (e.g. Van der Burg and Van Gorp 2005; Höffe 1993; Hillerbrand 2010; Ross and Athanasoulis 2012; Nihlen Fahlquist 2015) who address the balancing of moral values. However, we wish to add that phronesis also concerns the balancing of moral *and* epistemic values, acknowledging their entanglement.

Note that our suggestion for the realization of the third practice turn is not to replace standard decision approaches (like risk minimization or precautionary approaches) with virtue ethics. Rather, we seek to supplement these approaches with a virtue theory that has phronesis, as an intellectual virtue focusing on moral issues, at the center. The problems and complexities we face when dealing with risky technologies call for renewed attention to the ancient concept of phronesis.

9.5 Training Practical Wisdom in/After the Third Practice Turn

The previous sections have argued that risk is a thick concept that necessitates normative and descriptive considerations right from the start, calling for a third practice turn in risk analysis. We claimed that contemporary approaches to risk seem to add uncertainty as a kind of secondary complication to the ethical analysis when ethical principles are adopted for certain decision situations. We said that contemporary approaches to risk thus fail to take the entanglement of risk's normative and descriptive aspects seriously. And we contended that the intellectual virtue of phronesis, as an intellectual judgment that is devoted to the morally good, provides one path towards an integral risk analysis that considers the ethical aspects as well as the descriptive aspects of the uncertainties involved right from the start.

Introducing virtues provides a paradigm shift in contemporary risk analysis. In the latter, the focus is mostly on the right actions or outcomes, not on the attitudes or virtues of the deciding moral subjects. These subjects can be the policymakers who control the market launch of a new technology or the engineers who develop a new product, but also the members of the public who should be included in decision-making. In this section, we want to discuss two ways in which phronesis as the intellectual virtue for dealing with risky technologies may be developed and cultivated. These two ways involve engagement with emotions (Sect. 9.5.1) and with art (Sect. 9.5.2). We will argue that emotions and art allow phronesis to be sensitized and made aware of the normative aspects of risk.

9.5.1 Risk Emotions as a Way to Train the Phronesis

A dominant approach in social science from the last decades argues that one should involve the public in decision-making about risky technologies in order to contribute to democratic justification. A challenge for such an approach is that it requires laypeople to be willing and able to inform themselves about the complexities of risky technologies. Another challenge is often seen in the fact that people respond emotionally to risk. Debates about technological risks related to issues such as nuclear energy, climate change, and biotechnology frequently give rise to intense emotions. The dominant approaches in the literature to risk and emotions follow Dual Process Theory (Kahneman 2011). They consider emotions to be in conflict with rationality and thus a threat to rational decision-making (Sunstein 2005). Some scholars think that even though emotions are supposedly irrational, they should be respected for democratic reasons (Loewenstein et al. 2001).

However, while rationalist and quantitative approaches certainly have their virtues (as outlined in the preceding sections), they cannot adequately capture the *ethical* aspects of risk. While various risk scholars from philosophy and the social sciences have argued that we should indeed include ethical or 'qualitative' considerations in risk evaluations, they have not foreseen a role for emotions in this. They have either focused on rational capacities, explicitly denied a role for emotions in such processes (Sunstein 2005), or at most seen emotions as a heuristic that is highly prone to be biased and that should be corrected by reason (Loewenstein 2001; Slovic 2000; Slovic et al. 2004). Yet based on a cognitive theory of emotions, one can argue that we should understand risk-emotions as a form of practical rationality and as a potential source of moral wisdom (Roeser 2006, 2009, 2010a). Taking emotions seriously is crucial in debates about technological risks because emotions can point to what matters morally. Emotions such as sympathy, empathy, compassion, enthusiasm, and indignation can draw our attention to ethical aspects of risk – such as autonomy, justice, fairness, and equity – that are not included in quantitative approaches to risk (Roeser 2006, 2007, 2010a, b). For example, feelings of responsibility, shame, and guilt can lead one to awareness of moral obligations to future generations that might suffer from our use of non-sustainable energy sources. Feelings of indignation and anger may make us aware of violations of autonomy in the case of involuntary risk impositions.

Furthermore, emotions are especially well-suited to context-specific moral judgments; they provide us with more fine-grained sensitivity to concrete saliences than do rational judgments that are devoid of emotions (Roeser 2011). Emotions should therefore be a key ingredient in moral deliberation about risks (Roeser 2012; Taebi et al. 2012, cf. work on political emotions in general by e.g. Hall 2005; Kingston 2011; Staiger et al. 2010; Nussbaum 2013), and we want to argue that they can be used to train and strengthen the dianoetic virtue of phronesis. This approach offers a fruitful alternative to current theories that either neglect emotions and concomitant moral values or hold that emotions put an end to rational debate. Even approaches to participatory risk assessment do not explicitly involve emotions (Roeser and Pesch forthcoming). In contrast, the alternative approach presented here stresses that emotions should be a *starting point* for moral discussion and reflection about risk (Roeser 2012; Nihlén Fahlquist and Roeser 2015; Roeser and Pesch 2016). They are the key feature in identifying the potential ethical implications of future technologies, which is one of the tasks of phronesis defined in Sect. 9.4. This insight provides important input for the third practice turn in the study of risk by offering a sharper focus on moral aspects of risk – aspects that get excluded from technocratic approaches that ignore emotions and concomitant moral values.

However, emotions can also be biased. If emotional biases rest on scientific misunderstandings, then they need to be corrected by science (Sunstein 2010). But risk-emotions can also be morally biased, as by egoistic concerns. Such moral biases need to be critically examined by moral reflection. Here emotions themselves can play an important role (Lacewing 2005). This holds especially true for cognitive, moral emotions such as shame, guilt, and feelings of responsibility, with which we can critically reflect upon our initial emotions and realize that we should revise them (Roeser 2010c, 2011). For example, feelings of responsibility can reveal that we have to contribute to societal projects and that we cannot reject them simply based on an egoistic perspective (Roeser 2010c). This is what we want to call ‘emotional-moral reflection.’

Hence, emotions can help us critically reflect on emotions and concomitant values; for example, other-regarding emotions can help us critically reflect on our more selfish emotions and take a more balanced stance. Emotions address risk in an integrated fashion, often not separating its ethical and descriptive aspects. Hence they do justice to the fact that risk is a thick concept, and they are sensitive to specific contextual aspects of concrete technologies – and to their scientific, societal and moral complexities. They can strengthen our capacity for making context-sensitive moral judgments ('phronesis') about risky technologies, for experts as well as for laypeople. The moral judgments of both experts and laypeople are an important ingredient in democratic decision-making about risky technologies, and emotions can contribute to the quality of these moral judgments by providing extra resources via sensitivity, compassion and imagination.

9.5.2 Art as a Way to Train the Phronesis

In the previous section, we argued that emotions can play an important role in training our phronesis to make us more aware of the normative aspects of risk that are entangled with its descriptive aspects, and that emotions can also help us in critical moral self-reflection. Yet it is often difficult to transcend one's own emotional-moral perspective. Emotions and moral views are shaped by the environment and culture in which people are raised, and they sometimes resist influences that challenge people's core values (Kahan 2012; Greene 2013; Haidt 2012). This may threaten to make public deliberation difficult. However, philosophers have argued that art can contribute to emotional-moral reflection (e.g. Nussbaum 2001) and to politics (Kompridis 2014). Slovic and Slovic (2015) argue that art can make our experiences meaningful by means of emotions. Indeed, art can allow us to transcend our given emotional-moral perspective by appealing to our imagination and compassion.

The last decades have seen a more critical societal stance where many technological developments are concerned. At the same time, certain contemporary artists have become more and more interested in these developments. These artists' creations, which we will call 'techno art,' explicitly reflect on and engage with risky technologies.

Currently, there are numerous producers of techno art. Adam Zaretsky experiments with biotechnology, creating, for example, zebrafish with two heads. Eduardo Kac has developed a fluorescent rabbit, Alba, via genetic modification. The Tissue Culture and Art Project has created so-called 'victimless leather,' a delicate miniature coat grown via tissue engineering. These 'bioart' projects give rise to ethical and legal questions about the use and limits of biotechnology.

Relatedly, there is also art that engages with the ethical issues surrounding radioactive waste disposal. William Verstraeten designed an award-winning building for the Dutch nuclear waste facility COVRA and made artworks for its interior. His work reflects the ethical ambiguity of radioactive waste. The orange paint of the COVRA building, so bright that it is almost fluorescent, symbolizes both the danger and the unavoidability of radioactive waste storage. Over the next decades, the building will be repainted in progressively duller shades of orange, symbolizing the decay of radioactive waste.

Another important field of techno art is climate art, which engages with climate change and with technologies, such as climate engineering and geoengineering, that seek to combat it. Artist David Buckland runs a large interdisciplinary project, Cape Farewell, that aims to raise climate change awareness. Boo Chapple has created an interactive project that plays with geoengineers’ suggestion that climate change can be counteracted if we shield the earth under a white layer to reflect sunlight away. Chapple asked people to wear reflecting white hats and to deliberate on the desirability and impact of such technologies.

The interactions among technology, science, and art are rich and complex: technology nurtures fresh artistic developments by offering new methods and techniques, but artists also critically reflect on technology and inspire others to do the same. They use their work to explore ambiguities, paradoxes, and puzzlements. This allows them to contribute to public dialogues on science and technology, not merely by following technological developments but also by critically scrutinizing them (cf. Zwijnenberg 2009; Reichle 2009). Images and narratives provided by artists and writers can have a strong impact on people’s emotions and risk perceptions, thus shaping cultural discourse. According to Gessert (2003), artists can create awareness. They take potentially morally problematic aspects of technology to and over the limit, exploring their ultimate implications in works of art that go beyond what is at that point common practice. Zwijnenberg (2009, p. 19) argues that bioart can cue ethical reflection on scientific and technological developments more directly than theoretical texts can. This is because such works confront us with the possible implications of existing and future technologies concretely, materially, via an aesthetic experience.

Hence, many contemporary artists engage with risky technologies that give rise to emotions, highlighting the moral dilemmas that might be linked to these technologies and making potential contributions to moral and political debates. Art can provide new insights into the ethical aspects of risky technologies even before they have been developed. Art assists us in picturing different technological futures, which is essential for dealing with risky technologies morally – a central task of phronesis, as depicted in Sect. 9.4. In this way, art can contribute to the development of phronesis by facilitating explicit ethical reflection on the normative aspects of risk from the design phase onwards. Art can shape the third practice turn in the study of risk.

9.6 Conclusion

We started our chapter by identifying three practice turns in the social and philosophical study of technology: the first was the rise of social studies of engineering and technology, the second was the ‘empirical turn’ in philosophy of technology, and the third is the recent imperative to incorporate ethical values into a prospective ethical assessment. We then argued that these three developments can be seen in the social and philosophical study of risk as well: the first practice turn appears in the emergence of empirical decision theory and risk perception studies, the second turn surfaces in the social and philosophical study of risk as a societal and ethical phenomenon, and we argued for a third turn that addresses ethical considerations right from the start. We contended that this third turn is a necessary next step and a natural continuation of the first two turns. Only in this way can the interwoven nature of the descriptive and normative aspects of risk be properly taken into account. We elucidated this claim by mentioning the normative assumptions underlying risk analysis and expected-utility approaches, and by highlighting their morally controversial assumptions in particular. We also underscored the normative assumptions that are inevitably involved in the seemingly descriptive analysis of technology. We then showed that classical approaches, such as quantitative risk analysis and the precautionary principle, fail to take into account both descriptive and normative aspects in an integrated fashion; we argued that a practical judgment modeled after the ancient notion of phronesis may be a suitable supplement to decision theoretic approaches, allowing them to overcome this shortcoming. We concluded the chapter by proposing to strengthen or educate this virtue, which helps to assess and critically reflect on ethical aspects of risk. We said that emotions and art can play an important role in this, by providing us with uniquely rich insights into moral values in the context of risk. They can permit us to engage imaginatively with future scenarios and with other people, sensitizing us to important moral issues. We believe that this set of approaches, i.e. the broad concept of phronesis, supplemented and aided by emotions and art, provides promising avenues for a third practice turn in the study of risk. It offers rich, context-sensitive, imaginative methods for an integrated and empirically informed ethical evaluation of risk that includes descriptive as well as ethical-normative aspects from the outset.

Acknowledgements We would like to thank Peter Kroes and Maarten Franssen for their extremely helpful comments on earlier drafts on this chapter, and Veronica Alfano for proofreading and editing our text.

References

Asveld, L., & Roeser, S. (Eds.). (2009). *The ethics of technological risk*. London: Routledge/ Earthscan.

Beck, U. (1986). *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. Suhrkamp: Frankfurt am Main.

Biddle, J., & Winsberg, E. (2010). Value judgements and the estimation of uncertainty in climate modeling. In P. D. Magnus & J. Busch (Eds.), *New waves in philosophy of science* (pp. 172–197). New York: Palgrave Macmillan.

Brook, B. W., Edney, K., Hillerbrand, R., Karlsson, R., & Symons, J. (forthcoming). Energy research within the UNFCCC: A proposal to guard against ongoing climate-deadlock. *Climate Policy*. doi: 10.1080/14693062.2015.1037820.

Chang, R. (1997). Introduction. In R. Chang (Ed.), *Incommensurability, incomparability, and practical reason* (pp. 1–34). Cambridge, MA: Harvard University Press.

Churchman, C. W. (1948). *Theory of experimental inference*. New York: Macmillan.

Churchman, C. W. (1956). Science and decision making. *Philosophy of Science, 22*, 247–249.

Decker, M. (2013). Technikfolgen. In A. Grunwald (Ed.), *Handbuch Technikethik* (pp. 33–38). Stuttgart: Metzler.

Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science, 67*, 559–579.

Douglas, M., & Wildavsky, A. (1982). *Risk and culture. An essay on the selection of technological and environmental dangers*. Berkeley: University of California Press.

Foucault, M. (1991). Governmentality. In G. Burchell, C. Gordon, & P. Miller (Eds.), *The Foucault effect* (pp. 87–104). Chicago: The University of Chicago Press.

Franssen, M. (2006). The normativity of artefacts. *Studies in History and Philosophy of Science Part A, 37*(1), 42–57.

Frisch, M. (2015). Predictivism and old evidence: A critical look at climate model tuning. *European Journal for Philosophy of Science, 5*(2), 171–190.

Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L., & Keeney, R. L. (1981). *Acceptable risk* (p. 185). Cambridge: Cambridge University Press.

Gessert, G., (2003). Notes on the art of plant breeding. In *L’Art Biotech catalogue, exhibition catalog*, Nantes: Le Lieu Unique, 47

Giddens, A. (1990). *Consequences of modernity*. Cambridge: Polity Press.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. London: Viking.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Intuitive judgment: Heuristics and biases*. Cambridge: Cambridge University Press.

Greene, J. (2013). *Moral tribes*. New York: Penguin.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage Books.

Hall, C. A. (2005). *The trouble with passion: political theory beyond the reign of reason*. New York: Routledge.

Heilbroner, R. L. (1967). DO machines make history? *Technology & Culture, 8*, 335–345.

Hillerbrand, R. (2010). Unintended consequences and risky technologies. A virtue ethical approach to the moral problems caused by genetic engineering. In D. Pavlich (Ed.), *Environmental justice and global citizenship* (pp. 167–183). Amsterdam: Rodopi.

Hillerbrand, R. (2011). Von Risikoabschätzungen zum „guten Leben“, *Aus Politik und Zeitgeschichte, Themenheft Ende des Atomzeitalters? 61. Jahrgang, 46–47*(14), 42–48.

Hillerband, R. (2015). The role of nuclear energy in the future energy landscape: Energy scenarios, nuclear energy and sustainability. In B. Taebi & S. Roeser (Eds.), *The ethics of nuclear energy: Risk, justice, and democracy in the post-Fukushima era* (pp. 231–249). Cambridge: Cambridge University Press.

Höffe, O. (1993). *Moral als Preis der Moderne: Ein Versuch über Wissenschaft, Technik und Umwelt*. Frankfurt a.M: Suhrkamp.

Janssen, P. H. M., Petersen, A. C., van der Sluijs, J. P., Risbey, J. S., & Ravetz, J. R. (2005). A guidance for assessing and communicating uncertainties. *Water Science and Technology, 52*(6), 125–131.

Jasanoff, S. (1993). Bridging the two cultures of risk analysis. *Risk Analysis, 13*(2), 123–129.

Jeffrey, R. (1954). Valuation and acceptance of scientific hypotheses. *Philosophy of Science, 22*, 237–246.

Jonas, H. (1979). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Frankfurt a.M: Insel Verlag.

Kahan, D. (2012). Cultural cognition as a conception of the cultural theory of risk. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of risk theory: Epistemology, decision theory, ethics, and social implications of risk* (pp. 725–760). London: Springer.

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*(4157), 1124–1131.

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis, 1*(1), 11–27.

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis, 8*, 177–187.

Kirchin, S. (Ed.). (2013). *Thick concepts*. Oxford: Oxford University Press.

Kingston, R. (2011). *Public passion: rethinking the grounds for political justice*. Kingston: McGill-Queen's Press.

Kompridis, N. (Ed.). (2014). *The aesthetic turn in political thought*. London: Bloomsbury Academic.

Kroes, P., & Meijers, A. (Eds.). (2000). *The empirical turn in the philosophy of technology*. Amsterdam: JAI.

Kroes, P., & Meijers, A. (2006). The dual nature of technical artefacts. *Studies in History and Philosophy of Science Part A, 37*(1), 1–4.

Künneke, R., Mehos, D. C., Hillerbrand, R., & Hemmes, K. (2015). Understanding values embedded in offshore wind energy systems: Toward a purposeful institutional and technological design. *Environmental Science & Policy, 53*, 118–12.

Lacewing, M. (2005). Emotional self-awareness and ethical deliberation. *Ratio, 18*, 65–81.

Little, M. (2000). Moral generalities revisited. In B. Hooker & M. Little (Eds.), *Moral particularism* (pp. 276–304). Oxford: Clarendon Press.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welche, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267–286.

Luntley, M. (2003). Ethics in the face of uncertainty: Judgement not rules. *Business Ethics A European Review, 12*(4), 325–333.

Maga, D. (2015). *A methodology to assess the contribution of biorefineries to sustainable bio-based economy*. Oberhausen: Karl Maria Laufen.

McDowell, J. (1981). Non-cognitivism and rule-following. In S. Holtzman & C. Leich (Eds.), *Wittgenstein: To follow a rule* (pp. 141–162). London: Routledge & Kegan Paul.

Moore, G. E. (1988 [1903]). *Principia Ethica*. Cambridge: Cambridge University Press.

Möller, N. (2012). The concepts of risk and safety. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of risk theory: epistemology, decision theory, ethics, and social implications of risk* (pp. 55–85). Dordrecht: Springer.

Nihlén Fahlquist, J. (2015). Responsibility as a virtue and the problem of many hands. In I. van de Poel, L. Royakkers, & S. D. Zwart (Eds.), *Moral responsibility and the problem of many hands* (pp. 187–208). London: Routledge.

Nihlén Fahlquist, J., & Roeser, S. (2015). Nuclear energy, responsible risk communication and moral emotions: A three level framework. *Journal of Risk Research, 18*(3), 333–346.

Nussbaum, M. (2001). *Upheavals of thought*. Cambridge: Cambridge University Press.

Nussbaum, M. (2013). *Political emotions: why love matters for justice*. Cambridge, MA: The Belknap Press of Harvard University Press.

O'Malley, P. (1999). Governmentality and the risk society. *Economy and Society, 28*(1), 138–148.

Peinke, J., Böttcher, F., & Barth, S. (2004). Anomalous statistics in turbulence, financial markets and other complex systems. *Annals of Physics, 13*(7–8), 450–460.

Raynor, S., & Cantor, R. (1987). How fair is safe enough? The cultural approach to societal technology choice. *Risk Analysis, 7*(1), 3–9.

Reichle, I. (2009). *Art in the age of technoscience: Genetic engineering, robotics, and artificial life in contemporary art*. Vienna: Springer.

Reitinger, C., Dumke, M., Barosevcic, M., & Hillerbrand, R. (2011). A conceptual framework for impact assessment within SLCA. *The International Journal of Life Cycle Assessment, 16*(4), 380–388.

Reitinger, C., Kopriwa, N., Penner, H., Piana, L., Dumke, M., Fayyaz, S., Hillerbrand, R., & Pfennig, A. (2012). Integration sozialer Aspekte bei der Prozessbewertung. In M. Decker, A. Grunwald, & M. Knapp (Eds.), *Der Systemblick auf Innovation. Technikfolgenabschätzung in der Technikgestaltung, Tagungsband NTA 4* (pp. 419–423). Berlin: Sigma.

Roeser, S. (2006). The role of emotions in judging the moral acceptability of risks. *Safety Science, 44*, 689–700.

Roeser, S. (2007). Ethical intuitions about risks. *Safety Science Monitor, 11*, 1–30.

Roeser, S. (2009). The relation between cognition and affect in moral judgments about risk. In L. Asveld & S. Roeser (Eds.), *The ethics of technological risk* (pp. 182–201). London: Routledge/Earthscan.

Roeser, S. (Ed.). (2010a). *Emotions and risky technologies*. Dordrecht: Springer.

Roeser, S. (2010b). Intuitions, emotions and gut feelings in decisions about risks: Towards a different interpretation of “neuroethics”. *The Journal of Risk Research, 13*, 175–190.

Roeser, S. (2010c). Emotional reflection about risks. In S. Roeser (Ed.), *Emotions and risky technologies* (pp. 231–244). Dordrecht: Springer.

Roeser, S. (2011). *Moral emotions and intuitions*. Basingstoke: Palgrave Macmillan.

Roeser, S. (2012). Risk communication, public engagement, and climate change: A role for emotions. *Risk Analysis, 32*, 1033–1040.

Roeser, S., Fahlquist, J. N., Hillerbrand, R. (2015). Risk. In H. ten Have (ed.), *Encyclopedia of Global Bioethics* (pp. 1–10). Springer, online first.

Roeser, S., & Pesch, U. (2016). An emotional deliberation approach to risk. *Science, Technology & Human Values, 41*, 274–297.

Roeser, S., Hillerbrand, R., Sandin, P., & Peterson, M. (Eds.). (2012). *Handbook of risk theory: Epistemology, decision theory, ethics, and social implications of risk*. London: Springer.

Ross, A., & Athanasoulis, N. (2012). Risk and virtue ethics. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of risk theory: Epistemology, decision theory, ethics, and social implications of risk* (pp. 833–856). London: Springer.

Rudner, R. (1953). The scientist qua scientist makes value judgments. *Philosophy of Science, 20*(1), 1–6.

Shrader-Frechette, K. (1991). *Risk and rationality*. Berkeley: University of California Press.

Slovic, P. (1999). Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield. *Risk Analysis, 19*, 689–701.

Slovic, P. (2000). *The perception of risk*. London: Earthscan.

Slovic, S., & Slovic, P. (Eds.). (2015). *Numbers and nerves: Information, emotion, and meaning in a world of data*. Corvallis: Oregon State University Press.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24*(2), 311–322.

Soler, L., Zwart, S., Lynch, M., & Israel-Jost, V. (Eds.). (2014). *Science after the practice turn in the philosophy, history, and social studies of science*. London: Routledge.

Staiger, J., Cvetkovich, A., & Reynolds, A. (2010). *Political emotions*. London: Routledge.

Starr, C. (1969). Social benefit versus technological risk. *Science, 165*, 1232–1238.

Steger, F., & Hillerbrand, R. (2013). *Praxisfelder angewandter Ethik. Ethische Orientierung in Medizin, Politik, Technik und Wirtschaft*. Münster: Mentis.

Sunstein, C. R. (2005). *Laws of fear*. Cambridge: Cambridge University Press.

Sunstein, C. R. (2010). Moral heuristics and risk. In S. Roeser (Ed.), *Emotions and risky technologies* (pp. 3–16). Dordrecht: Springer.

Taebi, B., & Kloosterman, J. L. (2008). To recycle or not to recycle? An intergenerational approach to nuclear fuel cycles. *Science and Engineering Ethics, 14*(2), 177–200.

Taebi, B., Roeser, S., & Van de Poel, I. (2012). The ethics of nuclear power: social experiments, intergenerational justice, and emotions. *Energy Policy, 51*, 202–206.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.

van de Poel, I., & Royakkers, L. (2011). *Ethics, technology and engineering*. Oxford: Wiley-Blackwell.

van den Hoven, J., Vermaas, P. E., & van de Poel, I. (Eds.). (2015). *Handbook of ethics, values, and technological design sources, theory, values and application domains*. Dordrecht: Springer.

Van der Burg, S., & Van Gorp, A. (2005). Understanding moral responsibility in the design of trailers. *Science and Engineering Ethics, 11*, 235–256.

Williams, B. (2006). *Ethics and the limits of philosophy*. London: Routledge.

Zwijnenberg, R. (2009). Preface. In I. Reichle (Ed.), *Art in the age of technoscience: Genetic engineering, robotics, and artificial life in contemporary art* (pp. xii–xxix). Vienna: Springer.

Rafaela Hillerbrand Ph.D. in Theoretical Physics, University of Münster, Ph.D. in Philosophy, University of Erlangen. Currently Professor of Philosophy of Science and Technology, Institute for Technology Assessment and Systems Analysis (ITAS) & Institut für Philosophie, KIT-The Research University in the Helmholtz Association, Germany and head of the institute for philosophy at KIT. Research interests: modeling in science and engineering, computer simulations, ethics of energy supply and demand, sustainability, design and modeling for values, capability approach.

Sabine Roeser is Professor of Ethics at TU Delft (distinguished Antoni van Leeuwenhoek Professor) and head of the Ethics and Philosophy of Technology Section. She is head of the Human Research Ethics Committee (HREC) of TU Delft, and is member of numerous governmental advisory boards on risky technologies. She plays an active role in the Dutch public debate via interviews and public lectures. Roeser has been working at Delft University of Technology since September 2001. From September 2010 through August 2013 she also held a part-time professorship at Twente University (Socrates Chair). Roeser holds degrees in philosophy (Ph.D., M.A. cum laude), political science (M.A.) and fine arts (B.A.). She has published numerous books and articles. Her research covers theoretical, foundational topics concerning the nature of moral knowledge, intuitions, emotions and evaluative aspects of risk, but also urgent and hotly debated public issues on which her theoretical research can shed new light, such as nuclear energy, climate change and public health issues.

Chapter 10
The Policy Turn in the Philosophy of Technology

Adam Briggle

Abstract The empirical turn has been framed far too much in terms of *what* philosophers say and not to *whom* they speak. I apply the logic of the empirical turn to the very philosophers who carry its banner. I argue that once we look at them through their own lens, we discover that the empirical turn is not such a revolutionary thing after all. It is a turn within the disciplinary model of knowledge production. In other words, its own material culture and political economy look just the same as so-called classical philosophy of technology. In contrast, I sketch out what a policy turn looks like, which is a turn toward a new model of philosophical research, one that begins with real-world problems as they are debated in public and cashes out its value in real-time with a variety of stakeholders. I conclude by sketching some of the main ramifications of taking a policy turn in the philosophy of technology.

Keywords Relevance • Socially-engaged philosophy • Impact

10.1 Tasting Technology

When one stops talking in abstract ways about capital “T” Technology and begins discussing the specifics of various technologies in their manifold contexts, one has taken the empirical turn. In their edited volume that did much to inaugurate this movement, Peter Kroes and Anthonie Meijers (2000) define the empirical turn as one of “opening the black box” of technology to uncover the detailed, variegated richness therein. They talk about the need for an “empirically adequate description” of technologies – and it is always “technologies” in what might just as well be called the pluralist turn.

A. Briggle (✉)
University of North Texas, Denton, TX, USA
e-mail: adam.briggle@unt.edu

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_10

It's as if the old, classical philosophy of technology was a horrible social boor imbibing all sorts of fine wines at a tasting party and pronouncing each one to be "Wine." If you've tasted one, you've tasted them all. I take the empirical turn to be a kind of refinement of our senses – a scholarly *arbiter elegantia* if you will. What it's aiming at is a more subtle appreciation of the many wines on hand. None of it is just "Wine." This one is buttery with an oaken nose and an apricot finish. That one has an angular punch. That one has a cigar-box complexion with chewy tannins and an earthy grip. The other one there is jammy and opulent. In the same way, that is not just "the Internet" but social media sites, Facebook, twitter, and the thousand-fold ways of interacting with their thousand-fold features.

It's good for one's own soul to be a sensitive connoisseur of the world. It will heighten the pleasure one garners from experience. But is it doing anyone else any good? I use this kind of decadent metaphor on purpose. I want to suggest that there is something self-indulgent about the philosophy of technology whether in its classical or empirical form. I mean, the "turn" that has been taken is one *within* the same party. It is advocated because it will be good for us – us philosophers, that is. It will improve our philosophical reflections. But is that the end goal or is it on its way to some larger good? Are we "tasting" technologies simply to be better tasters or to put our refined sensibilities to some use?

Kroes and Meijers hint that the empirical turn is not just about good taste. In one brief comment, they say that such a turn is necessary in order "to be taken seriously in present-day discussions about technology." There is no indication of what these discussions are, but presumably they are ones being had by non-philosophers – maybe engineers, policy-makers, and various stakeholders. I think that "being taken seriously" by such people should be a major and explicit goal for philosophers of technology. And that would require a turn outside of the self-absorbed activity of tasting-to-be-better-tasters.

Many philosophers of technology will make a passing comment about how they could be "practically relevant" in all sorts of controversial issues. Yet that's almost always just a promissory note – perhaps well-intentioned, but rarely well-conceived let alone cashed out with work conducted in the midst of those controversial issues.

The bias is that the *real* philosophical work is to be had in clearing the ground and making empirically adequate descriptions of technology. Just how that is going to be of practical relevance is not clear. Maybe we aspire to help others to appreciate and critique technology just as well as we can. But there's always so much work to be done in getting the descriptions right. We tarry at the party discussing things with our fellow adepts. Yes, we are empirically grounded. But we are grounded in a wine-tasting party that is itself floating free…out of the sight of those people who might take us seriously.

10.2 Turning the Empirical Turn Inward

I think the reason for this is that the empirical turn hasn't been sufficiently applied to itself. Philosophers haven't examined in rich detail the one technology that is most close at hand, namely, the knowledge-technology of the academic discipline. They haven't done an empirical analysis of themselves as peculiar kinds of artifacts.

Note the title of Kroes and Meijers' introduction to the agenda-setting volume: "A Discipline in Search of Its Identity." The implication here is that philosophy of technology is, *naturally*, an academic discipline. But that's pure neutrality. It's the blank canvass or the empty box upon which or into which one then adds "an identity."

But this is the same kind of black-boxing that the empirical turn is supposed to shun. Philosophy of Technology is treated in just the same transcendentalist way Heidegger treated Technology (cf. Verbeek 2005). Heidegger is accused of treating Technology as the condition for the possibility of the modern worldview, rather than itself an object (rather, *many* objects) for inspection. I am accusing the empirical turn of doing much the same thing. The discipline of philosophy is treated as simply the condition for the possibility of thinking, rather than itself an occasion for thinking.

I'm not just hoisting the empirical turn on its own petard. I am internalizing its logic; I'm turning the empirical turn inward. Philosophers of technology who are committed to the empirical turn should examine their own practices, their own artifacts, their own institutions.

What will they find if they stop inconsistently exempting themselves from their own empirical gaze? I think they might find that the supposedly momentous break brought about by the empirical turn starts to look like more of the same old stuff. An inward look at the techniques of philosophy itself would show that the differences between the classical and empirical schools are rather superficial. This search for an identity starts to look like a man rummaging around in a costume box, trying to decide which hat to wear. The empirical turn may be a new wardrobe, but it's not a new way of life.

The way of life here is the disciplinary one. It consists of philosophers producing articles and book chapters for fellow philosophers to read. Really, only a very few adepts could tell much difference between a "classical" and an "empirical" scholarly article. The material reality here is the same: computers in offices, reams of specialized journals, occasional trips to the library, and conference travel. It would all look like just so much scholasticism to anyone outside the discipline. I say it *would* look that way, because, of course, no one outside of the discipline is looking in the first place.

I think this is one way in which the empirical turn operates, namely, transforming apparent heterogeneity into sameness. For example, Heidegger's distinction between modern technology and *poesis* starts to wash away. After all, one can find a genuine relation to being even when immersed in digital media (or so the argument goes). Modern technology is poetic. Here too the classical/empirical divide dissolves – their political economies and empirical realities are the same.

10.3 The Policy Turn

Of course, there's another side of the coin here, because the empirical turn also (and most visibly) transforms apparent homogeneity into variation. It does so by calling our attention to differences and distinctions that had been overlooked. Heidegger's *Gestell* becomes a whole panoply of things to be conceptualized and categorized. It's very much like descending in an airplane or looking through a microscope – a brown blur becomes fields and farmhouses, a green blob becomes paramecia and mitochondria. So too, the empirical turn applied to philosophy of technology will call out important distinctions that had been glazed over. Slight bumps on a seemingly smooth surface would protrude as vital terrains in their own right.

So, when I look at this movement through its own lens, those little nods to "practical relevance" or "discussions" where we might be taken seriously begin to magnify. I see in them a crucial distinction that has been overlooked. I see the difference between practices of philosophy that reside squarely within the disciplinary model of knowledge production and those that depart from that model in various ways and to different degrees.

This distinction is the deceptively simple one between *what philosophers talk about* on one hand and *who philosophers talk to* on the other hand. The empirical turn has so far pertained primarily to the former issue, which is about the conceptual content of philosophy. The policy turn pertains to the latter issue.

In the spirit of the empirical turn, let's trace this terrain with more care. Because when we examine it closely at an even smaller scale with an eye toward making conceptual clarifications, a variety of pathways begin to appear.

For starters, note that the empirical quest to develop adequate accounts of technologies tends to require a certain kind of interdisciplinarity. There is, in other words, a natural impetus written into the heart of the empirical turn toward a questioning of and even a modest break from disciplinarity. Philosophers need to get out of the armchair for a while to actually *learn from* engineers or farmers or whoever works on the ground with the technologies of interest, because this is the only way to be sure one is garnering an empirically adequate understanding of them. And indeed, there are several examples of articles co-authored with philosophers and engineers.

However, this pathway is often circular, or perhaps centripetal. After learning about the technologies at hand, the philosopher returns to the disciplinary fold in order to talk to other philosophers. So, although there is an interdisciplinary moment in the process, the audience remains disciplinary, and intellectual value is cashed out in the traditional form of peer-reviewed scholarship. We philosophers use the engineers to enrich our understanding, but we offer nothing by way of return. Or perhaps we do, but there is no institutionalized or systematic way to collect evidence about the impacts we are having on other fields.

Let me contrast that with what I am calling the “policy turn,” which is something that my colleague Bob Frodeman has written the most about. Indeed, 10 years ago he advocated for a policy turn in environmental philosophy, writing: “A policy turn…means a shift from philosophers writing philosophy essays for other philosophers to doing interdisciplinary research and working on projects with public agencies, policy makers, and the private sector” (Frodeman 2006, p. 3).

The policy turn implies a kind of interdisciplinarity that is different from the centripetal version noted above. In that version, the spheres of knowledge production and knowledge use are still separate. The philosopher retires to his or her office to produce knowledge about a technology in a space (journal or book) that is far removed from the contexts where stakeholders engaged with that technology are making policies about it. If those policies are going to be influenced by the knowledge product, it would have to be through some indirect process of serendipitous diffusion.

By contrast, the policy turn seeks to reduce or even eliminate the gap between knowledge production and use, thereby making the impact of philosophy more direct. The philosopher works *in media res*, cashing out his or her contributions within the interstices of a real-world discussion about technology. The audience is composed of those non-philosophical stakeholders who care about the issues at hand. The material culture of philosophy changes as a result, as the philosopher works in multiple media chosen for their capacity to connect with the target audience. Time will be spent at city hall, in the lab, on the farm, at the factory…and that is where the philosophy will happen. These are not *Bestand* to be harvested for raw materials and taken back to the ivory tower where the “real” philosophical work can be done. The philosophizing is in the interacting.

Let me sketch a pair of generic examples of my criticism, one drawn from recent issues of the two leading journals in the field, namely *Techne* and *Philosophy & Technology*. In the first instance, an article about cognitive enhancement technologies develops a critical analysis of the ethical and political implications of emerging trends. Yet there is no discussion of how this analysis is to be implemented – by which actors in which institutions or through which rule changes, etc. In the second case, reflection on the meaning of digital technologies for education turns up fascinating insights. Yet here too there are no remarks on how these ideas might be brought to bear on education by administrators, teachers, parents, or students.

Of course, in some sense these are not fair criticisms, because the artifact that is the scholarly journal article is not designed to serve the kind of function that I have in mind. If you are aiming to influence physicians, nurses, educators, parents, etc., then you would not write an article in a philosophy journal. So, my point is not to criticize individual contributors to these journals. Rather, my point is to question the place of peer-reviewed scholarship in the axiology of academic life. Why are peer-reviewed publications the gold standard for hiring, tenure, and promotion?

I very much want philosophers of technology “to be taken seriously” in present-day discussions about climate change, de-extinction, energy policy, food policy, and a host of other tech-intensive issues. But having an empirically adequate description of technologies won't be enough if those descriptions are hermetically sealed in the insular correspondences of a set of specialists. This means that philosophers of technology need to reconsider what counts as doing the real work of philosophy. It is time to stop asking whether we are sufficiently empirical and start asking whether we are sufficiently relevant.

10.4 A Twenty-First Century Philosophy

By a "policy turn," I mean a turn away from a peer-centric model of scholarship to one that is directly engaged in an ongoing technical controversy with various stakeholders. One could imagine a variety of such contexts and, thus, might imagine an "education" turn, a "military" turn, a "medical" turn, a "farming" turn, etc. in the philosophy of technology. I use the term "policy," however, broadly to encompass any of these topical areas. In each instance, there would be stakeholders involved in decision-making processes bringing a variety of perspectives to bear in a quest to define and secure special and common interests.

Thus, when I speak of a policy turn, I imagine philosophers working in dynamic partnerships with a range of people that might include parents, community groups, elected policy makers, bureaucrats, scientists, engineers, farmers, architects, manufacturers, and whoever else has a stake in a decision that involves technoscientific elements.

It is often asked: What kind of "engagement" does a philosopher have in such situations? That is to say, what does the philosopher have to offer *qua philosopher* to the ongoing tumult of policy processes? I think at the most general level of abstraction the philosopher offers abilities to identify, clarify, and critique conceptual and normative dimensions of the issue at hand. Philosophers are good at challenging claims to expertise and authority, uncovering hidden value judgments and assumptions, recognizing and critiquing various arguments and framing devices, offering creative alternatives, and posing fundamental questions that are often overlooked. These contributions can be made in a variety of modes, including that of an honest broker of dialogue and that of an issue advocate championing a specific policy direction.

To delve further into the nuances of a policy-oriented philosophy would require much more space than I wish to use here. The intent of this essay is more polemical than programmatic. I want to provoke reactions rather than lay out the nuts and bolts of a policy turn in the philosophy of technology. Much more detail can be found in the recent book *Socrates Tenured*: *The Institutions of 21st Century Philosophy* (Frodeman and Briggle 2016). And you can find a book-length case study of my own policy-oriented philosophical work in *A Field Philosopher's Guide to Fracking* (Briggle 2015).

However, I can say a bit about what kind of transformation the philosophy of technology would need to undergo in order to realize a policy turn. At its core, this change requires an expansion of the audience for the field's research. Certain key stakeholders, especially policy makers, need to become the explicit focus of attention rather than just fellow disciplinary peers. There are more or less radical ways to do this.

The less radical way is to tack on some additional activity at the back end of the normal peer-reviewed publication process. So, a philosopher can study an issue and publish a journal article about it, but rather than hope that somehow it might get taken up by society, the philosopher can actively send the article to important stakeholders making decisions about the issue. The next step would then be to set up a meeting with those groups to see if there is a way to implement the ideas in the article.

The more radical form of transformation entails leaving the disciplinary model of scholarship altogether – or at least not *beginning* with it. In this case, the philosopher would engage in something like what anthropologists call "participant observation," which entails beginning with immersion in a policy issue. The philosopher would get to know the stakeholders involved, learn about the issue, and begin to insinuate philosophical contributions wherever they may be of some use. The idea is that the philosophical dimensions of the issue will be interstitial in nature – popping up here and there – such that the philosopher needs to be "on the ground" and "in the mix" to make contributions in a timely and effective manner.

This shift in audience requires a change in the way we evaluate research and think about what 'real' and 'excellent' philosophy mean. The disciplinary model defines excellence in terms of peer-review, which is cashed out in the traditional form of bibliometrics (number of publications, citations, h-index, etc.). These other models of scholarship, especially ones that break with disciplinarity entirely, will require alternative metrics and accounts of excellence. Most importantly, those metrics and accounts will have to factor in the assessments made by the target audiences that are not one's philosophical peers – e.g., the policy makers involved in the issue at hand.

I want to conclude by highlighting some of the other main ramifications of a policy turn in the philosophy of technology. I do this both to provide a bit more insight about where I think this turn would take us as well as to head off some common misunderstandings.

First, in keeping with the empirical turn, I am a pluralist about models of philosophizing. I am not opposed to disciplinary practices. I am only opposed to their monopoly. The policy turn is consistent with disciplinary work. It just decenters disciplinarity as the only legitimate model of research.

Second, I am aware that some philosophers of technology practice something like the policy turn already. They do so with technology assessment organizations and with other stakeholder groups in a variety of contexts. Yet these occasions are all too rare and there is hardly any self-reflexive talk about what best practices for such work might look like. No one is training the next generation. Policy-oriented or socially-engaged research remains inchoate and inarticulate. The policy turn now is mostly the one-off adventures of daring individuals, rather than a sustained and institutionalized effort (see Briggle et al. 2015).

Third, the policy turn could save the life of philosophy. We live in an audit or accountability culture that increasingly demands some measure of “returns on investment” when it comes to research. The disciplinary model offers only a hand waving faith in the serendipity of ideas. The policy turn, by contrast, would allow philosophers to make, measure, and narrate more direct impacts on society. Again, not everyone would need to do such work. But a solid number of philosophers doing so would help provide herd immunity for all of philosophy in an age of increased scrutiny.

Fourth, the policy turn implies a new pedagogy for graduate students in philosophy. It frames philosophy as a practice, with more focus on helping students spot the philosophical moments residing within other disciplines and everyday life. Traditional courses and materials would be supplemented with activities designed to train a skill for becoming effective philosophical participants in the interstices of real-world issues. Internships would become the norm. This kind of education would also open up new pathways for employment of philosophers.

Fifth, disciplinary philosophy has fetishized rigor. Of course we should make good arguments; but theoretical excellence is a plural concept, relative to the temporal, economic, and rhetorical needs of a given audience. The policy turn calls for a reconceptualization of what counts as “real” scholarship. Indeed, in doing the kind of immersive work briefly described above, the philosopher will wrestle with a different kind of ‘hard’ or ‘rigorous’ work – one that entails political and rhetorical considerations of communicating to different audiences in addition to the difficult work of hammering out sound arguments and conceptual clarifications. This different kind of rigor deserves greater scrutiny by philosophers. In other words, the implementation of philosophical ideas ought to become a philosophical project in its own right. Philosophy needs a research program on the impacts of philosophy – which should in turn prompt the development of a philosophy of impact.

In summary, the philosophy of technology community needs to confront several issues that are simultaneously theoretical and pragmatic. This would constitute a reflexive research agenda that will be a necessary compliment to the philosophical work out in the field. We need to develop curricula to train next generation scholars to practice engaged work with diverse stakeholders. We need to reconceive our notions of rigor and how the epistemic elements are always intertwined with more arbitrary social and institutional factors. We need to revise tenure and promotion standards to reward broader social impacts on the same level as intellectual impacts on disciplinary peers. And, as part of this, we need to develop a broad set of evaluation standards for defining excellence outside of the disciplinary model of scholarship and for measuring the broader social impacts of philosophers.

References

Briggle, A. (2015). *A field philosopher’s guide to fracking*. New York: W.W. Norton.

Briggle, A., Frodeman, R., Barr, K. (2015). Achieving escape velocity: Breaking free from the impact failure of applied philosophy. *London School of Economics Impact Blog*.

Frodeman, R. (2006). The policy turn in environmental ethics. *Environmental Ethics, 28*(1), 3–20.

Frodeman, R., & Briggle, A. (2016). *Socrates tenured: The institutions of 21st century philosophy*. Lanham: Rowman and Littlefield.
Kroes, P., & Meijers, A. (Eds.). (2000). *The empirical turn in the philosophy of technology*. Amsterdam: Elsevier.
Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency, and design*. Chicago: University of Chicago Press.

Adam Briggle PhD in Environmental Studies from the University of Colorado, Boulder. Associate Professor in the Department of Philosophy and Religion at the University of North Texas. Member of the Executive Editorial Board for *Philosophy & Technology*. Research interests: the intersections of science, technology, ethics, and politics. Latest books: A Field Philosopher's Guide to Fracking (2015) and Socrates Tenured: The Institutions of Twenty-First Century Philosophy (2016 with Robert Frodeman).

Chapter 11
A Coherentist View on the Relation Between Social Acceptance and Moral Acceptability of Technology

Ibo van de Poel

Abstract According to the empirical turn, we should take empirical facts into account in asking and answering philosophical, including ethical, questions about technology. In this chapter, the implications of the empirical turn for the ethics of technology are explored by investigating the relation between social acceptance (an empirical fact) and moral acceptability (an ethical judgement) of a technology. After discussing how acceptance is often problematically framed as a constraint to overcome, a preliminary analysis of the notions of acceptance and acceptability is offered. Next, the idea of a logical gap between acceptance and acceptability is explained. Although the gap is accepted, it is also argued that the distinction between acceptance and acceptability does not exactly map on the descriptive/normative distinction and that both notions are maybe best seen as thick concepts. Next, it is shown how a coherentist account of ethics, in particular John Rawls' model of wide reflective equilibrium can account for the relation between acceptance and acceptability.

Keywords Acceptance • Acceptability • Technology • Wide reflective equilibrium • Thick concepts • Naturalistic fallacy

11.1 Introduction

The introduction of new technologies into society sometimes leads to social resistance. Examples are the building of nuclear plants, chemical factories and of wind mills which have, on occasion, led to public resistance or stakeholder protests. Sometimes the resistance concerns not the building or construction of a concrete artefact or plant, but rather the technology as such, sometimes even before actual

---

I. van de Poel (✉)
Delft University of Technology, Delft, The Netherlands
e-mail: i.r.vandepoel@tudelft.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_11

new products have been developed or are on the market. This is for example the case with protests against such technologies as biotechnology and synthetic biology, and to a lesser extent nanotechnology and geo-engineering. In still other cases, there are no loud or outspoken protests, but people simply do not use or adopt a new technology or product, which may also be seen as a case of non-acceptance.

Lack of acceptance is often seen as a problem by innovative companies that develop new technologies but also by universities and policy makers who help to develop, or who financially support new technologies. Non-acceptance is then perceived by them as due to, or a sign of unwillingness or irrationality of the public and therefore as a barrier to overcome. Conversely, social groups and non-governmental organizations are sometimes quick to point out that the lack of acceptance is not based on unwillingness or irrationality bur rather signals the moral unacceptability of these technologies.

Both reactions are in my view too simple. Although, lack of acceptance of a technology may point at the moral unacceptability of that technology, we cannot simply conclude from non-acceptance that a technology is also morally unacceptable. Lack of acceptance may have other reasons including (but certainly not restricted to) a lack of information, unjustified fears, selfishness or a limited ability or willingness to scrutinize the relevant moral reasons. Conversely, non-acceptance is not necessarily irrational or based on wrong reasons, and therefore a barrier to overcome. It may also point at good reasons not to accept a technology and hence be a sign of the moral unacceptability of a technology.

The main question of this paper is how we should conceive of the relation between acceptance and acceptability of a technology. To answer this question, I start with sketching what I consider a problematic framing of the issue, namely one in which lack of acceptance is merely seen as a barrier to overcome rather than as a potential sign of the non-acceptability of a technology. After sketching this problematic framing, I turn to the main question by first presenting a preliminary analysis of the notions of acceptance and acceptability. Next I will argue that (non)acceptance does not entail (non)acceptability of a technology. After pointing out this logical gap between acceptance and acceptability, I will argue that the terms are nevertheless related, first by arguing that both are probably best seen as thick concepts and then by showing how a coherentist notion of ethics can explain the relation between the two concepts.

11.2 A Dominant Framing: Acceptance as Problem to Be Overcome

In public debates, but also in policy documents, acceptance is often framed as a problem to be overcome. Take for example the debate on sustainable energy technologies, like wind mills. Wind turbines have raised public resistance, for example because of the noise they produce, the nuisance due to the moving shadows of the wind mill blades and the effects on the landscape. Typically, proponents of wind mills recognize the existence of such objections, but they frame them as barriers to overcome.

Once resistance or non-acceptance is framed as a barrier to overcome, the natural next step is to think of measures to take the barrier away. Typical measures that are often proposed are better informing people or financial compensation, but one might also think of more sophisticated approaches. For example, in the case of wind mills, ownership of farmers might contribute to better acceptance. Another measure to increase acceptance might be to create local cooperatives that have a say in the use of the wind mills and share in the revenues.

Although measures to increase acceptance are not necessarily objectionable, it is important to realize that the framing of acceptance as a barrier is in danger of leading to a one-sided approach because once acceptance is solely seen as a barrier, it can no longer be a source of serious social and moral concern about a technology. It is just a constraint to overcome. Moreover, this constraint is then often treated similarly to other constraints such as technical and legal ones.

Basically, there are two approaches to constraints when trying to realize a technology. First, some constraints simply have to be accepted. For example, it seems sensible to accept the physical constrain that we cannot make a perpetual motion machine when we try to develop new energy technologies. Other constrains may, however, be overcome. A material might not be strong enough currently for use in a wind turbine. It might, however, be possible through technological development and innovation to improve it so that this technical constraint is overcome. Likewise, it may be possible to change legal constraints, for instance one that forbids the use of unmanned vehicles on the road.

If social acceptance is construed as a constraint, then it is to be expected that it will be treated somewhat similar to other constraints. This means that social non-acceptance either has to be accepted or attempts may be made to overcome it. In cases the constraint is accepted, one might try to work around it, for example it might be decided to site a wind mill (or a nuclear plant) at another location than initially intended due to a lack of public acceptance. In other cases, one might want to overcome the constraint by clever ‘social engineering’; I already mentioned examples like institutional arrangements of (shared) ownership.

What is most problematic about framing acceptance as a constraint to be satisfied is that it favors an instrumental attitude towards this constraint: whatever the reasons we may have been for requiring acceptance of a technology, accept this requirement (constraint) as a given fact and find ways (means, instruments) to satisfy this constraint. If acceptance is just a constraint, and non-acceptance a barrier to overcome, any successful attempt to satisfy that constraint appears to be a step forward. What counts is the successful implementation of a technology in society. However, this way of looking at things runs the risk of hiding two important issues. First of all, not all means to achieve acceptance of a technology may be morally acceptable; thus apart from the practical question whether a particular means will be effective and efficient in realizing acceptance, the moral question whether the end justifies these means will have to be addressed. Secondly, and more importantly, the lack of acceptance may be the result of rejecting a technology for *moral* reasons. The de facto non-acceptance of a technology may be the result of the fact that the technology is considered to be morally unacceptable. In that case we are dealing with yet another *morally normative* issue, namely about the moral and social *acceptability* of a technology. By framing acceptance as a constraint to overcome, we run the risk of losing sight of the moral issues that might be involved in (non-)acceptance. In sum, then, the framing of acceptance as a constraint is problematic because it uncritically assumes both the moral acceptability of certain goals (introducing a technology in society) and of certain means (to overcome the constraint).

It should be noted that overlooking these moral issues may also be ineffective in improving the actual acceptance of a technology. For example, if people feel that they are 'bribed' by a financial compensation to overcome the negative consequences for them of for example the siting of a hazardous facility, then that may decrease rather than increase their willingness to accept that facility (Zaal et al. 2014). The point here, however, is not that it is ineffective to see acceptance solely as a barrier to implementation or to neglect the moral issues; rather the point is that it leads to a morally unacceptable reduction of the normative and moral issues at stake.

The aim in this paper is to develop an account of the notions of acceptance and acceptability that allows us to avoid the problems sketched above. This requires an account that on the one hand acknowledges the difference and logical gap between acceptance and acceptability, a task that I take on in the next two sections, but that at the same time is able to account for the relation between the notions of acceptance and acceptability, which I will address in the Sects. 11.5 and 11.6.

11.3 A Preliminary Analysis of the Notions of Acceptance and Acceptability

I will not make an attempt to come up with precise definitions of the notions of acceptance and acceptability. Both notions may be defined and operationalized in many different ways depending on the specific context at hand. For the current purpose, it is sufficient to point out that at the core there is an important distinction between acceptance and acceptability. The core distinction is that acceptance is primarily a descriptive notion while acceptability is primarily a normative notion.

With making a distinction between descriptive and normative notions, I do not wish to imply that the descriptive and normative can always be neatly separated. In practical judgments, they may be interwoven in such a way that they can hardly be separated (see the discussion on thick concepts below). This, however, does not preclude an analytic distinction between the descriptive and the normative. In my opinion any useful or plausible distinction between acceptance and acceptability should recognize that acceptance is mainly a descriptive notion, and acceptability is mainly a normative notion.

The analytic distinction between the descriptive and normative I have in mind here by and large follows the conventional philosophical distinction. The descriptive refers to what is, was or will be the case or is possibly the case. In other words, it refers to a description of actual, past, future or possible state-of-affairs of the world. Typically the way we have access to the descriptive is through observation, including introspection, and (other) forms of scientific investigation.

The normative, roughly, refers to what is good or desirable and what we ought to do. Often a distinction is made within the normative between the evaluative and prescriptive. The evaluative refers to certain normative evaluations, such as "this is a good person", or "this is a safe car"; while the prescriptive refers to what we ought to do or not to do (it refers to the deontic or to the right); for example, "you ought not to kill". Another distinction within the domain of the normative is the one between the practically normative and the morally normative. The practically normative centers around instrumental goodness ("this is a good hammer" and "turn on the wheel for changing direction"), whereas the morally normative concerns moral goodness ("this is a good person"). The study of instrumental goodness and the design of good instruments is primarily the domain of engineering, which heavily relies on the descriptive sciences. By contrast, the study of the morally normative does not (primarily) rely on the sciences or observation (although according to some it may rely on intuition or practical judgments) but rather it is the subject of (philosophical) ethics and moral philosophy.

If one subscribes to the idea that acceptance is basically a descriptive while acceptability is basically a normative notion, these notions may be further explicated along the following lines.

Acceptance as a (mainly) descriptive notion refers to certain states-of-affairs (that might be or not be the case). These states of affairs are typically those in which users, the government, the public or other stakeholders accept a technology. This of course still leaves open a number of issues that have to be resolved in order to operationalize the notion of acceptance.

For example, what does it mean to say that users accept a technology? Do users accept a technology when they use it so that we can perhaps look upon use rates as a proxy for user acceptance? Or is it possible that someone uses a technology because he or she has no alternative in which case it may be seriously questioned whether use implies acceptance? Perhaps we should measure acceptance by asking people and doing a survey? Another issue is that we should probably allow for degrees of acceptance, as it is typically the case that some people accept a technology while others do not. But acceptance also does not seem a matter of just counting numbers; individuals may also accept a technology to a smaller or lesser degree; and it might be that the acceptance of some people is more important than that of others (certainly if acceptance is seen as barrier to implementation of a technology, the acceptance of some actors is more important than that of others). Nevertheless, in all these cases acceptance seems to refer to certain states of affairs, even if one might debate what states of affairs are most relevant or important for establishing acceptance and even if some states of affairs may be more easily observable (and scientifically tractable) than others.

If we construe acceptability as primarily a normative notion, it refers to certain evaluations and/or prescriptions. As an evaluative concept, acceptability may be understood as expressing the evaluation that something is acceptable according to some normative standard. That standard might be moral norms and values, certain public values, a code of ethics or perhaps even moral standards that are laid down in the law; it might however also be a more implicit normative standard. What is important is that in such cases saying that a technology is acceptable not just means saying that it meets certain pre-given standards (which might be interpreted as the observation of a state of affairs) but implies certain judgments. Among others, it implies the judgment that the standards by which the evaluation is pursued are the right standards to judge this case and that there are no other normative considerations or standards that are left out of the evaluation. Just as statements about acceptance, judgments about acceptability are open to revision or correction, but for a different reason. Statements about acceptance are open to revision because of changes in interpretation of what has been *observed*. Normative judgments are open to revision, for example since the normative standards against which something is judged may change because of moral reasons or concerns that were not previously taken into account.[1]

Acceptability as a normative concept not only has an evaluative but also a prescriptive or deontic dimension. When we call a technology acceptable, then that appears to imply that we ought to accept that technology or at least that it is desirable, or perhaps even less strongly, that is permissible to accept that technology. One can obviously debate about how strong the deontic operator (obligation, desirability, permissibility) is that is implied by calling a technology acceptable, like one can debate what states of affairs exactly need to apply to call a technology accepted. Still it seems hard to contest that acceptability does not only express certain normative evaluations, but has deontic ramifications as well.

11.4 The Logical Gap Between Acceptance and Acceptability

Starting from this rough delineation of the meanings of the terms acceptance and acceptability, I will now sketch what may be called the logical gap between both notions. The idea of such a logical gap is based on the assumption that acceptance is a descriptive notion and acceptability a normative notion. In the next section, I will argue that this assumption is too simple. Nevertheless I think it is important first to sketch the idea of a logical gap between the descriptive and the normative because it reveals an important argument why acceptance and acceptability do not entail

---

[1] Note that normative judgments about acceptability of a technology are also open to revision because of changes in the way the object of the normative judgment (the technology) is interpreted and described. This case comes closer to changes in statements about acceptance. For our purposes, however, this is not an interesting case; it may even be questioned whether in this case we are dealing with a normative judgment about the same technology.

each other, even if the distinction between acceptance and acceptability does not exactly map on the descriptive/normative distinction.

The idea of a logical gap between the descriptive and the normative goes back to philosophers like David Hume and G.E. Moore (who in other respects hold quite different positions in moral philosophy). Hume's law as it has been called by some (Hare, for instance) mainly states that it is impossible to derive prescriptive statements ('ought') from descriptive statements ('is'); the naturalistic fallacy of Moore states that evaluative statements (about good) cannot be reduced to statements about descriptive (naturalistic) properties.

The famous passage from Hume that is often quoted as establishing Hume's law reads as follows:

> In every system of morality, which I have hitherto met with, I have always remarked, that the author proceeds for some time in the ordinary ways of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when all of a sudden I am surprised to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, 'tis necessary that it should be observed and explained; and at the same time that a reason should be given; for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. (Hume 2000 [1739], 3.1.1.27)

It is a matter of controversy what Hume exactly says or claims here, but one important interpretation is that Hume claims that prescriptive statements cannot be derived from descriptive statements. Whether this is really what Hume intended to say is less important for the purpose at hand; what has become known as Hume's law clearly underlies the idea of a logical gap between the descriptive and the normative.

Unlike Hume, G.E. Moore is a moral realist who believes that moral statements are objectively true or false. However, despite the fact that his meta-ethical position is completely different from Hume's, he has put forward arguments that also imply a logical gap between the descriptive and the normative. This is generally known as the naturalistic fallacy: it says that good (a moral property) cannot be defined in naturalistic, i.e. descriptive terms. Moore, for example, believes that 'good' cannot be understood in terms of pleasures and pains (as Bentham did).

The main argument that Moore gives for the impossibility of defining moral terms in non-moral, descriptive terms is the open question argument (Moore 1903). The open question argument says, basically, that always when we state that something possesses some natural (descriptive) property that is allegedly equivalent to being good, we can meaningfully ask the question, but is it good? For example, when we say of an action that it is pleasurable (descriptive property), we can meaningfully ask the question, but is it good? This implies that 'being pleasurable' cannot be equivalent to being good, because otherwise the question, 'but is it good?' would be redundant.

The open question argument might indeed be directly applied to the distinction between acceptance and acceptability. According to the idea of a logical gap, always when we state that some technology is accepted, we can still meaningfully ask the question, but is it acceptable? Indeed it seems that we can always meaningfully ask this question. There is in fact a whole range of reasons why a technology may sometimes be accepted but still be unacceptable (or not be accepted while being acceptable).

One reason why a technology that is accepted may be unacceptable is that people may have accepted a technology on the basis of incomplete or wrong information; in such cases it might well be the case that the technology should not have been accepted, and is indeed unacceptable given the right information. Similarly, acceptance may be based on wrong (moral) reasons. Another possibility is that important parties may have had no voice in the acceptance of a technology, making a technology potentially unacceptable. Yet another possibility is that people have accepted something because they had no choice, there was no alternative; in such cases we might want to say that a technology is unacceptable despite it being accepted.

Note further that a reason why the open question argument applies to acceptability is that, as I noted above, the normative evaluation whether something is acceptable is not just an observation of a state of affairs but involves a moral or normative judgment. This suggests that even if we could spell out in complete detail the states of affairs that need to be the case to call something accepted, we can still meaningfully ask, whether being accepted amounts to being acceptable.

The idea of a logical gap between acceptance and acceptability, because one notion (acceptance) is descriptive and the other (acceptability) is normative, has a certain attractiveness. It is neat and simple. It puts questions of acceptance in the realm of observation and the sciences; while questions of acceptability are in the realm of ethics. It is also in line with the idea that we cannot conclude from the fact that a technology is accepted that it is also acceptable, or the other way around, that from the fact that something is acceptable, we cannot conclude that it is or will be accepted.

Still, I think it does not suffice to note that acceptance and acceptability do not entail each other. We also need an account of how both notions are related. Although such an account should acknowledge that the notions are different and do not logically entail each other, it should at the same time be able to give a positive account of the relation between both terms. In the next two sections, I try to develop such an account. I do so by first arguing that both notions can be seen as thick concepts that have descriptive as well as normative content. Next, I will show how a coherentist account of ethics, in particular John Rawls' model of wide reflective equilibrium, may account for the relation between acceptance and acceptability without assuming that the one entails the other.

11.5 Acceptance and Acceptability as Thick Concepts

Below, I will argue that the notions acceptance and acceptability are thick concepts, i.e. concepts that have both descriptive and normative content. This is not to say that we cannot distinguish acceptance and acceptability or that the one implies the other, but rather than their distinction does not neatly or completely match the descriptive/normative distinction.

Thick concepts are concepts that are descriptive and normative at the same time; their application is world-guided as well as action-guiding (e.g. Williams 1985). Thin normative concepts are concepts that are only action-guiding and have no descriptive content. Typical examples of thin normative concepts are very general normative concepts like good or just. There is, however, also a category of normative concepts that is at least in part world-guided and has normative content. A typical example is 'brave'. We only call, for example, a soldier brave if his behavior and character meet certain characteristics. In this sense, braveness requires some minimal descriptive content; whether it applies in a particular case is at least partly world-guided; not all kinds of behavior or all kinds of character traits count as braveness. At the same time, braveness is a normative concept as it implies a (positive) normative evaluation and it has certain deontic (and thus action-guiding) implications, at least for some kinds of people, like soldiers.

Although the idea of thick concepts is usually applied to normative concepts, it may be applied to descriptive concepts as well. That is to say, while there may be thin descriptive concepts that are purely descriptive, there may also be thick descriptive concepts that are partly normative or action-guiding in nature. In fact, I believe that acceptance is such a concept.

Earlier, we have seen that if we want to apply the notion of acceptance in a particular case, then it has to be operationalized which implies that a number of choices have to be made. For example: does mere use of a technology count as acceptance or does acceptance also require certain attitudes or intentions? Who should accept a technology in order to conclude that it is accepted? Does the acceptance of all relevant stakeholders weigh equally when we try to establish whether a technology is accepted or not? (and how do we delineate what are relevant stakeholders in the first place?)

So, to make the notion acceptance empirically tractable a number of methodological choices need to be made. These choices are normative, although not necessarily morally normative in nature. They may be based on for example epistemological concerns and values. For example, how to exactly interpret acceptance may depend on one's epistemological aims. Also what data are available or can be acquired may motivate choices to define acceptance in a particular way.

Apart from epistemological concerns, practical and moral normative concerns may also motivate adapting a particular definition and operationalization of acceptance. For example, if one aims to contribute to insights how to increase the acceptance of a technology, some definitions of acceptance are likely to be more suitable than others. For example, notions of acceptance that cannot be influenced (by for example policy measures) may in this case be less suitable than notions of acceptance that can.

Definitions of acceptance also have, at least implicitly, a moral component. For example, are only the direct users of a technology relevant in defining and operationalizing acceptance or also other stakeholders? One may approach this question as being about who influences the actual use and shaping of a technology and on that grounds decide not to include certain stakeholders; but one might also approach it from a moral point of view and argue that the acceptance of all who are potentially affected by a technology is relevant from a moral point of view, even if they do not influence actual technological implementation and development. It should be noted that even if such morally relevant choices are not explicitly made from a moral point of view, they imply an implicit moral choice and are in that sense morally normative even if they are not motivated by moral reasons.

We can thus conclude that acceptance is not purely descriptive because normative considerations, including moral ones, play, or at least may play, a role in how we exactly understand, define and operationalize acceptance. Even if such normative considerations do not explicitly play a role in operationalizing the notion of acceptance, the choices we made are still implicitly normative. Such normative choices, be it on epistemological, methodological, practical or moral grounds, are indeed indispensable if we want to make acceptance an empirically tractable notion. These (implicit or explicit) normative choices determine what states-of-affairs in the world are relevant for whether a technology is accepted or not.

Let us now turn to acceptability and see whether this largely normative notion has also descriptive content. In other words: is acceptability at least partly descriptive and dependent on acceptance? One argument why acceptability has a descriptive component may be that acceptance tracks moral reasons that are relevant for acceptability.

This is the case because it seems likely that for most people the question whether to accept a technology, or not, is motivated by normative, including moral, considerations. In other words, people's judgment when they accept (or do not accept) a technology are likely to express certain moral concerns and reasons. This is not to say that acceptance is always based on moral concerns or that it reliably tracks the normative reasons that make a technology acceptable or not. But even if acceptance does not imply acceptability, it seems likely that insofar as a moral judgement is implied in acceptance, it tracks at least some of the (potentially) morally relevant reasons that are also relevant for acceptability.

It could, however, be argued that this does not yet give acceptability descriptive content. On some meta-ethical accounts, moral reasons exist independent of what people think about them (e.g. Dancy 2002). In other words, acceptance may help to track normative reasons, but it is not necessary for the existence of such normative reasons nor can it create (new) normative reasons; acceptance is a heuristic device at best.

I think two objections are possible against the above argument that acceptability does not rely on acceptance and has no descriptive content at all. First, it may be argued that the idea that moral reasons can exist independent of what people think of them is mistaken. Instead, moral judgment seems to be always embodied, i.e. it is always judgment by people and moral reasons exist in the minds of people but not independent of hem. This does not mean that people's reasons are always or necessarily morally right; they may be mistaken, reflected upon and revised. What is morally right is not decided by what people believe, but moral beliefs and moral reasons always 'exist' as reasons in the minds of persons and are therefore at least in principle descriptively traceable. This would give moral reasons at least some descriptive content as they can be described as reasons that at least some people have or that motivate at least some people. Applied to acceptance and acceptability, it would mean that the moral reasons that are relevant for acceptability can be described as reasons that 'exist' in the mind of people when these people accept a technology or not.

A second possible argument why at least some form of acceptance is needed for acceptability goes as follows. An important moral value that seems relevant to the acceptability of a new technology is respect for autonomy. Respect for autonomy says that we should accept the morally autonomous and deliberate choices of people. In some contexts, respect for autonomy has been specified in terms of informed consent. For example in medical treatment, in clinical experiments or experiments with human subjects, informed consent is considered an important moral principle for deciding about the acceptability of a treatment or experiment.

Informed consent stipulates that a treatment or experiment, or more generally an intervention, can only be morally acceptable if people have given their consent. Valid informed consent is usually considered to be dependent on such conditions like whether the agent had all the relevant information, has understood that information and was able to make a voluntary choice (Beauchamp and Childress 2013, p. 124).

Although informed consent is not the same as (actual) acceptance of an intervention (it also depends on how acceptance is understood or defined), it seems clearly to have a descriptive component, i.e. whether people actually give their consent. So, if one believes that respect for autonomy, understood in terms of informed consent is an important moral principle in deciding about the acceptability of a technology, it seem that acceptability is at least in part dependent on some form of acceptance (i.e. that form of acceptance that is the result of informed consent).

This is not to say that acceptability solely depends on acceptance if one accepts respect for autonomy as an important moral principle. It has been pointed out that for example for clinical trials, informed consent is neither a necessary nor a sufficient condition for the acceptability of such experiments (Emanuel et al. 2000). Still, if respect for autonomy and informed consent are important principles that should be taken into account in judgments about the acceptability of a technology, it seems that we cannot simply ignore the fact that people accept or do not accept a technology in moral judgments about the acceptability of such a technology.

11.6 A Coherentist Account of the Relation Between Acceptance and Acceptability: Reflective Equilibrium

In the previous section, I argued that both acceptance and acceptability are thick concepts. Although this allows for a relation between both notions, it does of course not yet show that, and if so how, both notions are related. To do that we also need to point out that there is a relation between the content of both notions. I went some way in arguing for that possibility in the last two arguments in the previous section. I now will try to make it more plausible that both notions are also content-wise related by sketching how both notions relate to each other if we adopt a coherentist approach to ethics.

To do so, I will briefly sketch the reflective equilibrium model of John Rawls and apply it to the relation between acceptance and acceptability (Rawls 1999 [1971]; Rawls 2001). Rawls' model is based on a coherentist approach to ethics. According to coherentist accounts, there is no Archimedean point in ethics. That is to say there is no class of values, principles, norms or moral statements that is more fundamental or foundational than others and from which the validity or justification of all other moral judgments can be derived.

A coherentist approach thus deviates from for example a Kantian or utilitarian approach in which some moral principles are seen as more fundamental than others. In Kantian ethics, for example, this special foundational status applies to the categorical imperative and the 'good will'; in utilitarianism it applies to the value of human happiness and the principle that one should strive for the greatest happiness for the greatest number of people. Not only are such normative principles seen as given in foundational approaches, they are also conceived as the justification for more specific rules and judgments. For example, the judgment that one should not lie on a specific occasion derives its moral force in a Kantian approach from the categorical imperative that shows that one should not lie in general.

In foundational approaches, then, justification is ultimately based on a set of (moral) principles, values, norms or judgments that have a special status because they are founded in some special (external) way, be it that they are based on reason, on intuition or on human nature (or whatever other way). In coherentist accounts, there are no foundational moral principles, rules or values. Rather, justification is based on the coherence of the entire web of someone's moral principles, values, beliefs and judgments.

Coherentist approaches to ethics have been criticized because they lack any external grounding; according to some it might for example be possible to have a coherent racist ethical outlook. Another weakness of coherentist accounts is that as of yet, we lack a good operationalization of how to determine whether a web of moral principles, values, beliefs and judgements is more coherent than another.

Yet, coherentist accounts also have distinct advantages. One is that they recognize the fallible and provisionary nature of moral judgement. Rather than suggesting that moral judgments are given and infallible because they are based on some foundational and unchangeable moral values, they stress that moral judgments are always provisional.

What is more, coherentist approaches not only allow for the possibility of revision of moral judgments, they may also point out how such revision might occur. In a coherentist account moral revision is not based on a sudden struck of insight but rather it implies a revision of the entire web of someone's moral judgments, or at least a relevant part of it. Revision may for example begin with certain tensions between moral beliefs which may trigger changes in the entire web, so achieving new coherence (and new justification). This picture seems to strike a healthy balance between openness to moral revision and a certain conservatism when it comes to well-established moral norms and values (because revision in principle relates to the entire web of moral judgments, which will not be so easily overturned).

Rawls' model of (wide) reflective equilibrium may be seen as a specification of what a coherentist account of ethics could look like and how it allows for moral revision. Rawls' model distinguishes three layers of moral judgement: considered moral judgements, moral principles and background theories. These layers do not represent distinctions in moral force (or justification), but rather different levels of abstraction or generality. Background theories are usually the most general and abstract, moral principles are somewhat more specific and considered moral judgments are judgements about specific cases or situations.

Although revision may start at all three layers, it would seem that in the case of established moral belief systems, it will often start with new considered judgments. This is the case because considered judgments relate to specific situations or cases, and new situations may lead to new considered judgments. One important reason for this is that considered judgments are not only action-guiding, i.e. they do not just derive from moral theories and principles, but are also world-guided, i.e. they are also based on the facts of the specific situation or case.

If we apply this idea to our discussion about the relation between acceptance and acceptability of technology, we get something like the following picture. People's considered judgments about the acceptability of a specific technology (in a specific situation) may on the one hand been seen as 'facts' about the situation (i.e. the fact that people judge so and so). As we have seen, these facts are relevant for whether a technology is (actually) *accepted*. At the same time, such factual judgments feed into the judgement whether a technology is *acceptable*. The acceptability judgment, however, requires coherence between all three layers of the wide reflective equilibrium model. If someone's initial (considered) judgement is not coherent with his or her other moral beliefs, such coherence may be achieved either by revising the judgment or by revising other elements in his or her entire web of moral beliefs.

One way to understand the distinction between acceptance and acceptability in the light of Rawls' model of reflective equilibrium is to make use of the distinction between narrow and wide reflective equilibrium (Daniels 1979; Rawls 2001). In a narrow reflective equilibrium, we typically look for the moral principles and theories that fit our actual, given moral judgment best and we do not critically scrutinize our judgments about particular cases in the light of a range of moral theories and principles. Narrow reflective equilibrium is then more a descriptive mode and it reveals the moral ‘grammar’ that underlies actual moral judgments. De facto judgments that establish the *acceptance* of a technology seem to be based on such a narrow reflective equilibrium. In a wide reflective equilibrium, we more broadly reflect on our moral judgments, viewing them from a number of moral perspectives and principles, and striving for full coherence rather than just adapting our theories and principles to given judgments. Contrary to a narrow reflective equilibrium that is mainly descriptive, a wide reflective equilibrium *justifies* moral judgments. So, judgments about acceptability appear to require a wide reflective equilibrium whereas acceptance requires only a narrow equilibrium.

Especially in his later work, Rawls allows for the possibility that different people might come to different wide reflective equilibria (Rawls 2001). He considers this possible because people may have different world views, and even if the wide reflective equilibrium approach requires one to consider a range of moral perspectives and principles, it does not require accepting all these perspectives and principles and including them in the agent’s web of moral beliefs. In fact, it would seem impossible to include all possible moral perspectives and principles into one coherent web of moral beliefs. In this sense, even if a wide reflective equilibrium has a certain justifying power, it is not necessary that people come to the same wide reflective equilibrium and therefore to the same judgment for example on the acceptability of a technology.

Still, Rawls believes that for many public issues a so-called overlapping consensus is possible. An overlapping consensus is a moral belief (principle, judgment about a case, etc.) that is shared among the reflective equilibria of different people. Rawls believes that such an overlapping consensus is possible if not likely for those public issues for which people are willing to reason from public reason; i.e. on the basis of moral reasons that are relevant for the public domain and that accept what he calls reasonable pluralism, which include beliefs like that the state should not force people to accept certain nonpublic values and that it is possible to draw a boundary between public and nonpublic values.

The wide-reflective-equilibrium model and the notion of overlapping consensus can thus account for the (potential) difference between acceptance and acceptability, as well as for their relation. Acceptance can be understood as referring to people’s de facto judgements in a narrow reflective equilibrium, while acceptability refers to the judgment that people would have if they would achieve a wide reflective equilibrium. Moreover, as we have seen, people’s judgments on acceptability that are the result of achieving a wide reflective equilibrium by that person may still vary between persons. However, one might expect, or at least hope, that if people are willing to reason from public reason, an overlapping consensus will develop that represents a justified consensus on the acceptability of a technology.

One might then argue that according to a coherentist account, acceptance and acceptability coincide (or should coincide) in the ideal case. The reason for this is that in the case of full coherence (or wide reflective equilibrium), people’s (de facto) judgments about specific cases are in full coherence with the entire web of their moral beliefs, and are based on a wide reflective equilibrium. So their de facto judgments (acceptance) are justified due to the coherence with all their other moral beliefs and hence also express acceptability. In addition, in the ideal case that people reason from public reason, their judgements on acceptability will overlap, so representing an overlapping consensus on the moral acceptability of a technology.

In practice, however, this ideal situation is not very likely to be (completely) realized. First, there may and often will be differences between people's de facto judgments and the judgments they would have in wide reflective equilibrium. One reason for this is that full coherence is more a regulative ideal than a practical possibility. It does not seem unrealistic to assume that people's moral belief systems, just as their cognitive belief systems, always have some degree of lack of coherence. Although coherentist accounts of ethics postulate coherence as an ideal, it should be noted that an actual lack of coherence is not always bad (also not in the light of the ideal), as it can be a source of moral reflection, scrutiny and revision. In other words, lack of coherence may lead to moral reflection and that often seems a good thing.

Second, even if people's individual judgments are based on a wide reflective equilibrium, that is not yet a guarantee for an overlapping consensus with respect to the acceptability of a technology. Again, however, this lack of consensus need not necessarily be seen as only problematic, because it may also be a source of debate, argumentation and reflection. In other words, it would seem that reasonable people do not only form and adapt their wide reflective equilibrium as the result of intrapersonal deliberation but also as the result of deliberation with other persons. And again reasoning from public reason and the attainment of an overlapping consensus might be more a regulative ideal rather than a possibility that can be easily obtained in practice.

As stressed before, it is therefore important not to equate too easily acceptance and acceptability (even if they may coincide in the ideal case), because the distinction between the two invites moral reflection and debate.

11.7 Conclusions

In order to clarify the notions of acceptance and acceptability of technology and their relation I first presented some preliminary explications of both notions, after which I discussed the logical gap between the two notions. The idea of a logical gap has a certain attractiveness because it illuminates the differences between the two notions and it points to the danger of equating acceptance with acceptability. Still, I argued we also need to account for how both notions may be related. To this end, I first argued that both acceptance and acceptability are thick concepts that contain descriptive as well as normative elements. However, that does not mean that they cannot be distinguished. Next, by applying the wide-reflective-equilibrium model of John Rawls to the discussion, we have seen that we can conceive of acceptance as the result of a narrow reflective equilibrium, i.e. an equilibrium in which moral principles and background theories are adjusted to given considered judgments. Acceptability on the other hand would be the result of a wide reflective equilibrium that also critically scrutinizes considered judgements from a variety of moral theories and background principles. In contrast to a narrow reflective equilibrium, a wide reflective equilibrium also provides a moral justification and therefore has a strong moral force. It might be argued that in the ideal case of a wide reflective equilibrium with complete coherence, acceptance and acceptability coincide. This is, however, more a regulative ideal than a practical possibility. Moreover, even in wide reflective equilibrium, people might come to different moral judgments. They may still be able to reach an overlapping consensus if they reason from public reason as suggested by Rawls, but again this is merely a regulative ideal.

For debates about the acceptability of a technology the coherentist view implies that in judging the acceptability of technology we should take into account (rather than ignore) the stakeholders' acceptance (or non-acceptance) of a technology and that non-acceptance should not be simply treated as a barrier to overcome. However, the acceptance (or non-acceptance) by stakeholders need to be morally scrutinized, reflected upon and brought into coherence with other moral beliefs in a wide reflective equilibrium, before it can also be seen as an expression of acceptability. Moreover, people's moral judgments should be a source of debate rather than be seen as given and unchangeable.

More generally, a wide reflective equilibrium account, or for that matter other coherentist accounts, does not assume a neat distinction between normative and descriptive statements or judgments. In fact, at all three layers of the wide-reflective-equilibrium model, we find both normative and descriptive elements. Among the background theories we find not just moral theories but also descriptive theories, and often they will be a combination of both; like for example utilitarianism which is a moral theory but also a theory about human nature and about what makes people actually happy. Moral principles, the second layer in the model, will often be formulated with the help of thick normative concepts; that is to say that they also have some descriptive content. And as we have already seen before, considered moral judgments are both action-guiding and world-guided.

Although we may still be able to distinguish analytically between the descriptive and the normative, it may be difficult or impossible to unambiguously classify the elements that are in coherence in a wide reflective equilibrium model as either descriptive or normative. This also makes clear why we should take into account the empirical or descriptive in applied ethics, as also advocated in the empirical turn in the philosophy of technology, without assuming that normative or ethical questions now get empirical answers.

Acknowledgments This paper is the result of various conversations with Peter Kroes about this topic. I thank him wholeheartedly for these (and other) fruitful, inspiring and thought-provoking discussions. Earlier versions of this paper were presented at the symposium on "Social acceptance in energy through democratic and administrative procedures" held on November 6, 2014 in St. Gallen (Switzerland) and at the SPT (Society for Philosophy of Technology) conference, held from July 2–6, 2015 at Northeastern University in Shenyang (China). I also like to thank Maarten Franssen for comments on an earlier version of the paper. Any mistakes in the paper are fully my responsibility.

References

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics*. New York: Oxford University Press.

Dancy, J. (2002). *Practical reality*. Oxford: Oxford University Press.

Daniels, N. (1979). Wide reflective equilibrium and theory acceptance in ethics. *Journal of Philosophy, 76*(5), 256–282.

Emanuel, E. J., Wendler, D., & Grady, C. (2000). What makes clinical research ethical? *The Journal of the American Medical Association, 283*(20), 2701–2711.

Hume, D. (2000). [1739]. *A treatise of human nature.* Oxford: Oxford University Press.

Moore, G. E. (1903). *Principia ethica*. Cambridge: Cambridge University Press.

Rawls, J. (1999). [1971]. *A theory of justice.* Cambridge, MA: The Belknap Press of Harvard University Press.

Rawls, J. (2001). *Justice as fairness. A restatement*. Cambridge, MA: The Belknap Press of Harvard University Press.

Williams, B. (1985). *Ethics and the limits of philosophy*. Cambridge, MA: Harvard University Press.

Zaal, M. P., Terwel, B. W., ter Mors, E., & Daamen, D. D. L. (2014). Monetary compensation can increase public support for the siting of hazardous facilities. *Journal of Environmental Psychology, 37*, 21–30.

Ibo van de Poel is Anthoni van Leeuwenhoek Professor in Ethics and Technology at Delft University of Technology in the Netherlands. He has published on engineering ethics, the moral acceptability of technological risks, values and engineering design, moral responsibility in research networks, and ethics of new emerging technologies like nanotechnology. His current research focuses on new technologies as social experiments and conditions for morally responsible experimentation.

Chapter 12
Perovskite Philosophy: A Branch-Formation Model of Application-Oriented Science

Wybo Houkes

Abstract In this paper, I present a model of application-oriented science, to supplement existing work in science and technology studies on the re-orientation of scientific research. On this "branch-formation" model, research efforts may be guided by non-epistemic values without compromising their epistemic value: they may involve completion of mechanism representations that serve control over these mechanisms while also adding to our understanding of them. I illustrate this model with a case study from photovoltaic technology, involving the possible use of materials with the so-called 'perovskite' structure in dye-sensitized solar cells. The paper has three parts. The first argues how existing work on the increasing application-orientedness of scientific research can and must be supplemented with a perspective from the philosophy of science. The second presents the branch-formation model, which combines central ideas of the 'finalization-of-science' program of the Starnberg school with recent work in 'mechanistic' philosophy of science and in the philosophy of technology. The third part illustrates the branch-formation model with current developments in research on perovskite solar cells.

Keywords Basic and applied research • Mechanism • Mode 2 knowledge • Triple Helix

Scientific research does not take place in a vacuum. Even those who strongly believe in academic freedom and value-free science acknowledge that, in practice, research is partly driven by societal needs, as well as by more mundane external forces such as funding streams. Recently, however, many have claimed or voiced concerns that scientific research, especially at universities,[1] is *increasingly* focused on producing knowledge or other results that are directly relevant to other stakeholders: industry, governmental organisations, or society in general.

---

[1] In this paper, I mostly use 'research' or 'scientific research', and leave implicit that this research is traditionally (thought to be) done at universities.

W. Houkes (✉)
Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: W.N.Houkes@tue.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_12

There are several ways to characterise this change, depending on which aspect one wants to emphasise, and how it is assessed. Some note a shift in the scientific self-conception, towards a more “engineering” or “entrepreneurial” image (Daston and Galison 2007, p. 382ff; Etzkowitz 2002). Others regard it as an alarming move away from “pure” science, mainly under the influence of industry (Ziman 2000), as a transition, or as a revolution. In this paper, I label the change as a *re-orientation* towards *applications*. This label leaves it open whether researchers play an active role in redirecting their efforts, or mainly react to external incentives. It also highlights, among all interrelated changes, that many research efforts in especially the natural and life sciences are increasingly focused on devices and technologies. This focus means that the re-orientation of research is of interest to both philosophers of science and philosophers of technology, since it immediately concerns the interface or ‘trading zone’ between their areas of interest.

The actual interests of especially philosophers of science hardly reflect the ongoing changes in research practices. Researchers in science and technology studies and in innovation studies have, by contrast, been very active in studying how research is interwoven with industrial and societal concerns. In terms of publications in leading journals in the respective fields, there are more papers on the topic in some single issues of *Research Policy* than in entire volumes of *Philosophy of Science*.[2] Likewise, in interdisciplinary projects devoted to recent changes in scientific research, philosophers are seldom involved.

That philosophers of science and technology have not paid much attention to the re-orientation of scientific research does not, of course, mean that they should; nor that, as latecomers to the discussion, they will have much to add. I will argue in this paper that philosophers of science and technology *can* and *need to* contribute to studying, and in particular assessing, the application re-orientation of scientific research. Not to do so is an oversight, given the types of analysis that are currently available. Furthermore, I show that philosophy of science and philosophy of technology offer materials to construct a supplement to existing approaches, and I illustrate the supplement with an example of research on photovoltaic cells, in particular those containing materials with so-called ‘perovskite’ structure. This essay contributes to an empirical turn in the philosophy of science and technology: it uses and develops approaches in these sub-disciplines to improve our insight into an ongoing phenomenon – the application re-orientation of research – that has so far mostly been addressed in science and technology studies. Furthermore, it is partly empirical in method: after an initial sketch, the branch-formation model presented in the essay is developed in discussing a case study of ongoing application-oriented research.

Like perovskites, the paper consists of three parts. In Sect. 12.1, I review two of the most influential existing approaches to the re-orientation of scientific research: the ‘New Production of Knowledge’ programme and the ‘Triple Helix’ approach.

---

[2] Some publications in philosophy of science and technology are concerned with the application re-orientation, especially the growing role of commercial interests. Examples are Wilholt (2006), Radder (2010), and Irzik (2010).

I argue that these leave room and, more strongly, create a need for a supplement that looks at the contents of scientific research rather than its institutionalisation, and that allows an evaluation of the epistemic merits of application-oriented research. Philosophy of science and technology might offer the means to develop such a supplement. In Sect. 12.2, I outline one possible supplement, which I call the 'branch-formation' model of application-oriented science. It is built on the central tenets of the 'finalization-of-science' program, one of the few sustained efforts within the philosophy of science to reflect on the epistemic legitimacy of orienting scientific research towards broader, non-epistemic or 'societal' values. I show that the central ideas of this programme can be modified with elements of mechanistic philosophy of science and with work in the philosophy of technology on the design of artefacts. This leads to a model that represents research efforts as branches that develop representations of mechanisms, and that may share rudimentary versions of these representations – mechanism sketches or functional analyses – with other branches. Developments of branches and within branches may be guided by traditional scientific goals (i.e., creation of epistemic value[3]) through explaining and predicting phenomena; by external, societal goals such as increasing safety and sustainability through controlling phenomena in artificial devices; or by both types of values and goals. As an illustration of the latter, I reconstruct, in Sect. 12.3, the development of perovskite solar cells as a recent branch in photovoltaic research. In this area, the same mechanism sketch (a three-part functional analysis) is shared by several branches. Each branch seeks to complete the representation of the operating mechanism of a particular type of photovoltaic cell, in order to enhance the performance and durability of this cell, while minimizing the cost of production; thus combining scientific goals with societally or industrially relevant ones. Section 12.4 concludes.

12.1 Existing Approaches and the Need for a Supplement

In this section, I argue that philosophers of science and technology can and, to some extent, should contribute to studying the application re-orientation of scientific research. I do so by reviewing existing approaches and showing that these leave room and, more strongly, create a need for a supplement that may be developed by philosophers of science and technology.

Out of the many approaches to understand the ongoing re-orientation of scientific research,[4] I focus on the "New Production of Knowledge" (NPK; Gibbons et al. 1994)

---

[3] Throughout the paper, I sometimes use the distinction between (the creation of) epistemic and non-epistemic *values*, and sometimes that between (the pursuit of) internal/scientific and external/non-scientific (or 'societal') *goals*.

[4] Often called 'the science system' or 'systems' in the science-studies literature.

and the “Triple Helix” (3H; Etzkowitz and Leydesdorff 2000).[5] These are highly influential[6] and – despite their differences – point towards the same supplement. After reviewing both approaches, I give two reasons why such a supplement is possible or even necessary. First, the abovementioned frameworks are largely descriptive, or prescriptive only insofar as they propose policies to facilitate the changes that they describe. By contrast, philosophers could assess whether the phenomenon constitutes a *legitimate* change of scientific research, or one that interferes with its epistemic merits. Second, science-studies frameworks operate on a high level of aggregation: they focus on changes in science as a whole, as a knowledge-producing activity, and its place in society. By contrast, philosophers could focus, not on who produces knowledge in which organizational form and under which social conditions, but on *what* knowledge is produced for which more proximate purpose.

Both approaches resulted from research programs and thematic conferences, and harbour a variety of convergent views.[7] The outlines given here ignore most of this diversity for the sake of brevity.

The NPK approach revolves around a distinction between two ‘modes’ of knowledge production: the traditional ‘Mode 1’ and the new, emerging ‘Mode 2’. Mode 1 production primarily occurs within academic disciplines. These structure what counts as ‘good science’, both in determining what is a valuable contribution (i.e., criteria of relevance, originality, etc.), and in providing systems of quality control (in particular: review by scientific peers). In terms of physical location, Mode 1 production is ‘homogeneous’, because it is mostly a concern of research institutes such as traditional universities; in terms of interrelations with other societal groups or interests, it is ‘autonomous’, because it is not deeply or intrinsically concerned with the impact of the produced knowledge outside academic disciplines. By contrast, Mode 2 knowledge production is a heterogeneous, transdisciplinary affair. A larger variety of institutions – including government agencies, high-tech spinoff companies and consultancy firms – is involved in the production of Mode 2 knowledge, which irreducibly combines disciplinary perspectives and concepts. Furthermore, it is application-driven: knowledge production cannot be disengaged from its practical applications, because these structure what is ‘good science’ just as much as any disciplinary standard. Sensitivity to, for instance, the commercial or societal value of knowledge is an integral part of Mode 2 knowledge production, and such values enter into the quality control of its products – as universities in many countries have noted in the changing terms of assessment of their research.

---

[5] Hessels and Van Lente (2008) provide a more complete review of approaches, and focus on the New Production of Knowledge.

[6] The presentations of the NPK- and 3H approaches in this paper are based on Gibbons et al. (1994) and Etzkowitz and Leydesdorff (2000) respectively. These publications have been cited over 12,000 times and over 4500 times (Google Scholar, accessed August 2015).

[7] The NPK approach was originally presented in a single (multi-authored) book. The 3H approach is presented in several edited volumes and special issues, and not all contributions easily fit the same mould (some are by authors who later vehemently criticised the approach). Etzkowitz and Leydesdorff (2000), the introduction to one special issue devoted to the program, is taken as a guideline here.

Critics of the NPK approach have pointed out that its sharp dichotomy between two modes of knowledge production oversimplifies both the history of science and its contemporary diversity: some fields, such as pharmacology, might always have revolved around Mode 2 knowledge; conversely, it seems an overstatement that Mode 2 knowledge production is emerging in fields like particle physics. Relatedly, the various aspects of Modes 1 and 2 are only loosely interconnected, which undermines the explanatory potential of the approach. Since Mode 2 must, for instance, be both transdisciplinary and heterogeneous, possible cause-effect relations between these phenomena cannot be coherently formulated. Finally, it has been pointed out that the approach is biased: the focus on two ideal types, where one is explicitly associated with a 'traditional' way of doing science and the other is diametrically opposed to it, implicitly supports the transition to Mode 2 knowledge production – in all the aspects identified (Godin 1998).

The 3H approach can be partly understood as a response to the NPK approach. It does not posit a dichotomy between different ways of producing knowledge, with sets of diametrically opposed characteristics. Instead, it emphasises that, in any process of research or innovation, there is an interplay between industry, government and universities – the three strands of the helix. This interplay is in itself not a historical phenomenon, but any particular form that it takes is. Likewise, there may be different "resolutions of the relations among the institutional spheres of university, industry, and government" (Etzkowitz and Leydesdorff 2000, p. 110) in different countries, sectors of industry or academic disciplines. Strategies that allow mathematicians to cooperate fruitfully with Internet service providers in China might not apply to environmental scientists advising electricity companies in Germany. The flexibility of this approach concerns the relata as well as their relations. The historical forms of academia, industry and government partly co-determine each other. In a "central-planning" form of the helix, for instance, the state controls both industrial and knowledge production, and coordinates any relations between these activities; this interaction itself in part defines all three strands. Furthermore, the strands and their interrelations do not go through a random sequence of forms, but result from a "reflexive subdynamics of intentions, strategies, and projects" (ibid., p. 112), through which individuals and groups attempt to solve local problems, such as attainment of research funding or improvement of competitive power over commercial rivals.

In addition to this general framework, the 3H approach offers an account of the current "configuration" of the helix. In it, the three strands overlap; they partly take over each other's characteristic tasks or missions, and hybrid organisations emerge. Thus, we find government laboratories and spinoff companies on high-tech campuses, and universities are encouraged to take up a 'third mission' of economic development, besides teaching and research. The consequences that are outlined for scientific research closely resemble the aspects of Mode 2 knowledge production. Research is oriented towards particular contexts of applications, and needs to cross traditional disciplinary boundaries. Unlike in the nineteenth century, new areas of academic interest are not specializations within existing disciplines, but transdisciplinary fields like nanotechnology, artificial intelligence (Ahrweiler 1996) and, more contemporarily, data science. In many of these fields, research projects involve close collaboration with industrial partners, and occur in new "interstitial communities" and "interface locations" such as digital media laboratories and science parks, rather than at universities and research labs.

Criticisms of the 3H approach have no canonical form yet, which may be a result of its age and ongoing dynamics, as well as of ambiguities and unclarity in its central conceptions. Some critics (e.g., Shinn 2002) point out problems with the quasi-evolutionary terminology that is used in some of the seminal presentations. However, it is at best unclear whether this terminology is needed to apply the 3H approach to specific cases; many of the contributions to Triple-Helix volumes and special issues avoid such terminology. Secondly, it has been argued (e.g., Mirowski and Sent 2008) that the 3H approach is driven by the idea that universities should take up the third mission of creating economic value and should therefore become "entrepreneurial" (e.g., Etzkowitz 2002). In response, advocates of the approach emphasise that the Third-Mission idea is conceptually independent from the Triple-Helix idea. The latter indeed seems so general in its conception of industry-university-state relations that it can hardly entail the former. A third type of criticism is that, where the NPK approach may overemphasise differences, the 3H approach is overly focused on continuities – both in emphasising that scientific research has always been connected to industrial and state interests, and in downplaying any ongoing changes in its overall story of perpetual change. This feeds suspicions that its advocates, like those of the NPK approach, tacitly approve current changes in academia, but rather than appeal to revolutionary sentiments, they present a narrative of continuity that is geared towards alleviating any worries regarding these changes (e.g., Mirowski and Sent 2008).

The latter criticism has been formulated as an implicit commodification of knowledge, which might presuppose a neoliberal conception of the value of scientific research. It has been voiced against both approaches considered here (see also contributions to Radder 2010). Apart from shared preconceptions, it also reveals a shared level of analysis: both approaches focus on the overall organization of research, both internally (e.g., in transdisciplinary areas rather than in hierarchically organized fields of specialization) and externally (e.g., in interstitial communities). This organizational focus is especially clear in 3H work, which offers in-depth descriptions of how research in fields such as biomedical technology or nanoscience was (re-)organised in response to various external incentives, without discussing any of the content of research in these fields and how this might be different from more traditional research. Similarly, although the NPK approach is, at first glance, about *knowledge*, it does not feature studies of central research activities (such as theorizing, modelling, simulating, experimenting) that may be distinguished within the encompassing category of 'production' and that have been the specific focus of large bodies of work in the philosophy of science. Nor do the NPK and 3H approaches thematise how contexts of application may shape the content of scientific knowledge. Thus, the black box of scientific 'knowledge production' remains closed, in favour of focusing on the overall organization of universities, changes in patent law, etc.[8] Much of value may be found on this level of analysis, but work at this level cannot register significant changes in central practices and epistemic output and the disciplinary standards that adhere to them.

This ties in with a second, more contentious supplement. Seen from a philosophical perspective, the NPK and 3H approaches are, like much other work in science and technology studies, mostly descriptive: they signal changes in scientific research and its relations to industry, without assessing them. One example concerns the broader mechanisms of quality control signalled in the NPK-approach, which add considerations of economic value to disciplinary standards of excellence; this tendency is signalled, but not explicitly assessed as an unjustified imposition of standards external to science, a welcome correction to scientific self-centredness, or something between these extremes.[9] Such an explicit assessment might be expected from philosophy of science, given its traditional concerns.

Two remarks about the constraints on such an explicitly evaluative perspective are in order. First, evaluating the application re-orientation of science should not and need not involve a small set of standards, imposed by supposedly 'objective' philosopher-assessors. There is some hope that this can be avoided[10] by due attention to actual scientific research, following the example of the historical and practical turns in the philosophy of science. Second, and related, a credible assessment needs to take into account the diversity of scientific research. Here, the second supplementary role should initially build on the first: a micro-perspective on the knowledge production in a particular research area should facilitate assessment of the epistemic merits of work in this area. This assessment cannot be immediately translated to other areas, although it can be taken as a basis for further inquiry. A global evaluation of the re-orientation of scientific research is called for, but if this evaluation is to be sufficiently sensitive to actual practice, it can only proceed in a piecemeal fashion. This essay contributes a first piece, focusing on the first supplementary role of philosophy of science and technology. In the conclusion I will reflect on the limitations of the branch-formation model and on steps to develop a stronger, more encompassing evaluative perspective.

---

[8] As a case in point, all examples in Gibbons et al. (1994) are research *fields* rather than specific theories or other research products.

[9] The lack of explicit assessment of the shift to Mode-2 knowledge production and of the developments in the triple helix, makes it easier to accuse the NPK and 3H approaches of accepting or welcoming these changes.

[10] It is easier to identify these risks than to avoid them. Mirowski (2004) offers a revealing discussion of implicit biases in several influential programs in the philosophy of science.

12.2 The Branch-Formation Model: Finalization Meets Mechanisms

Most existing philosophical work limits the role of non-epistemic values in science to the choice of research topic (e.g., Kitcher 2001). Some (e.g., Douglas 2009) have argued that under certain conditions, such values may play a role in hypothesis testing, especially where there is inductive risk – roughly: a significant possibility of making errors with societally relevant consequences.[11]

Here, I develop another view of the role of non-epistemic values in scientific research. Central to the model of application-oriented science developed here is that research can encounter *branching points*, at which further development may, as far as epistemic values are concerned, legitimately depend on considerations such as societal relevance. Then, non-epistemic values may, in a way to be discussed in further detail, play a role in determining the *content* of scientific theories, as well as in the evaluation of their merits. This develops one central element of the NPK-approach, viz. that the 'context of application' co-determines scientific research; but it offers a more detailed and potentially normative version.

Construction of the branch-formation model proceeds by modifying the 'finalization-of-science' program, one effort by philosophers of science to reflect on the legitimacy of societal planning of scientific theorizing. This program ran for several years in the late 1970s and early 1980s at the Max Planck institute at Starnberg, (then West-)Germany. Some years after the seminal publications and a – to put it diplomatically – lively discussion of the program in German, its central contributions were translated and collected, together with a retrospective, in an edited volume (Böhme et al. 1983).

The finalization program, which is still occasionally mentioned in both the science-studies literature and more practice-oriented work in the philosophy of science, rests on three main ideas:

(a) Scientific disciplines go through three phases: a pre-paradigmatic phase in which basic theories need to be discovered; a paradigmatic phase in which basic theories are developed through 'normal science'; and a post-paradigmatic phase, which starts once theories are 'closed'.
(b) All disciplines ultimately end up with a closed theory, through research processes that are guided purely by goals internal to science.
(c) Once their theories are closed, scientific disciplines may be *finalised*, i.e., their further development may legitimately be guided by goals external to science.

The program thus combines ideas about the dynamics of science with a distinction between internal, scientific goals and external goals.[12] This leads to a clear

---

[11] See Biddle (2012) for an overview of these lines of work.

[12] From the context, it is clear that researchers in the finalization programme mainly had in mind the direction of scientific research by societal needs. Yet their central ideas are at least compatible with direction by other non-scientific purposes, such as the needs of industry or commercial interests.

specification of the conditions under which scientific research is legitimately employed in the pursuit of non-scientific purposes. Ignoring some variety and development within the programme, as well as amendments in response to criticisms, disciplines are said to be closed or mature once they have “theories which within their explanatory programme have formulated laws from which sufficiently precise and reliable predictions can be derived for the subject-matter the theory addresses” (Böhme et al. 1976, p. 309). Recurrent examples of such theories are mechanics, hydraulics, and hydrodynamics. Finalization means that goals external to science, in particular the creation of societal value, may be internalized by scientists to such an extent that these goals become the “guidelines” for specialization, differentiation and modification of the fundamental theory (ibid., pp. 311 and 315), for instance in disciplines such as chemical engineering, space research and agricultural chemistry.

Critical reception focused on the normative component of the programme and the supposed incompatibility with academic independence. However, its descriptive component seems more in need of modification. Judgements concerning the maturity, let alone completeness, of a field or theory are powerful rhetorical devices. Yet they are difficult to substantiate without begging important questions or making questionable assumptions about the structure of science. Relatedly, one might want to disengage the programme from the covering-law model of science that is implicitly assumed in the definition given above. This should lead to a descriptively more adequate model that retains the evaluative potential of the finalization programme.

The modifications proposed here take elements from mechanicist work in the philosophy of science. The first is that a central concern of scientific research is the discovery and accurate representation of the *mechanisms* that produce, underlie or maintain phenomena. Mechanisms are, to use the classical definition of this approach: “entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions” (Machamer et al. 2000, p. 3). Crucial for my purpose of modelling application-oriented research is that the mechanistic approach acknowledges that representation of mechanisms serves a variety of purposes – especially prediction, explanation and control of phenomena (Craver and Darden 2013, p. 6); and that representation is often incomplete: scientists typically start with *mechanism sketches*, in which only some of the entities, activities, organization or conditions are specified (e.g., Craver 2006). To take a stock example, discovery of the mechanism of protein synthesis involved developing a sketch that focused on the activities of DNA and RNA, and supplied additional entities, stages and conditions (Machamer et al. 2000, Sect. 6). Another way in which mechanisms might be sketched rather than supplied in full is in functional analyses, which indicate entities and activities by their systemic roles or functions rather than with structural descriptions. Such analyses can be developed into complete representations of mechanisms by replacing functional ‘placeholders’ with ‘bottom-out’ entities and activities (Piccinini and Craver 2011). To take an example from electrical engineering, a Wien bridge oscillator consists of four resistors and two capacitors, in a particular organisation (i.e., set of interrelations). A complete mechanistic explanation or design for (a phenomenon involving) such an oscillator needs to replace the functional terms ‘resistor’ and ‘capacitor’ with structural descriptions. Conversely, elements of more complete representations can be truncated or black-boxed to produce mechanism *schemas*, which can be re-completed or instantiated for other phenomena of interest.[13]

The dynamics of sketches, schemas and complete representations of mechanisms leaves room for minimally one type of scientific research that is effectively guided only by ‘internal’ goals – namely the construction of mechanism sketches. As long as too many of the entities and activities, organization and conditions involved in a phenomenon remain unspecified, scientific research has not produced sufficient information to be relevant for the pursuit of external goals, i.e., the creation of non-epistemic value. This holds in particular for mechanism sketches that cannot be straightforwardly reformulated as functional analyses: here, to put it metaphorically, it remains to determine the contours of the mechanistic map rather than to ‘merely’ fill white spots (i.e., black boxes). Yet many functional analyses also offer too little for industrial or societal purposes: as long as central entities and activities in – for instance – an oscillator remain unspecified, the costs, benefits, risks and rewards of producing and using it are up for grabs.[14] Thus, from an outsider perspective, there is no reason to guide researchers in their construction of mechanism sketches; it seems best to have them construct as many sketches as possible for the broadest conceivable variety of phenomena. The actual construction of some mechanism sketches may already be guided by non-epistemic as well as epistemic values – but *evaluating* them might as well be limited to the latter, since they offer little more than sketchy promises regarding the former.

Once a mechanism sketch has been produced, it may be developed or completed for different phenomena and for different purposes: representations of mechanisms may *branch out* from a shared sketch. Mechanistic philosophy of science has provided many examples, especially from the cognitive and life sciences, in which sketches and schemas need to be supplemented with additional elements of mechanisms, and in which structural descriptions replace functional analyses. As these examples show, sketches in themselves have little or no explanatory or predictive power; thus, scientific research cannot be said to be complete after sketches have been constructed. Moreover, since the representation of phenomena can change substantially while sketches are developed into more complete representations of mechanisms, even the ‘maturity’ of research that has led to sketches can be contested at any point. Sketches are never available as a *prêt à porter* commodity, or a finished product of ‘pure’ science, but rather form an initial product that needs further processing or even re-construction to be of use in scientific research.

---

[13] Craver and Darden (2013, Chs. 5–9) present a detailed and compelling case for the importance of mechanism schemas in science. I focus on mechanism sketches and only indicate possible roles of mechanism schemas in footnotes.

[14] In particular, mechanism sketches may fail to specify how well and under which range of circumstances an organized set of entities may be expect to perform certain activities. Thus, they may provide some superficial or ‘phenomenological’ understanding (Craver and Darden 2013, pp. 86–89), but are virtually useless for purposes of control.

Therefore, taking the construction of sketches as the ‘internal’ stage of scientific research, and the condition for later application-oriented research, avoids some of the problems of the maturity or completeness condition in the finalization programme.

Many of the branches that develop a mechanism sketch may be guided entirely by internal, scientific purposes; in fact, virtually all of the examples given in the literature concern such development of mechanism representations. It is noticed that mechanisms can also be represented to control phenomena, but the specific forms of developing mechanism sketches for this purpose have not received detailed attention. Here, earlier work in the philosophy of technology on the design of technical artefacts is useful.

Control of mechanisms may serve scientific purposes, for instance in setting up experiments or controlling for known confounding influences; and occasionally, control may serve no other purpose than demonstrating that something can be done. Yet many mechanisms are controlled because their termination conditions contribute to the realization of practical purposes, valued by individuals, organizations or society; and many mechanisms involve non-natural entities (such as conveyor belts and corkscrews) and/or activities intentionally undertaken by human agents (i.e., actions such as pushing and pulling). In many such cases, selection or production of the mechanisms can be understood as *design*; and fulfilling the starting conditions or engaging in some of the constitutive activities as *use*. Then, the entities in the mechanism can on one or more suitable levels of aggregation be ascribed technical functions.[15] The engine of a car is, for instance, the site of a carefully installed mechanism, which ought to contribute reliably to the fulfillment of transportation needs, as long as some agent – at the moment, a human being – starts the mechanism and engages in the activities known as ‘driving a car’. This is not a one-time, coincidental occurrence; the combustion mechanism is the result of purposeful design, and triggering of the mechanism is likely to be an instantiation of equally purposeful and recurrent use. Hence, the components of a car, like its engine, and sub-components like its carburetors have technical functions. And although it is most natural to think about the processes internal to technical devices as mechanisms, a mechanistic analysis can be extended to larger-scale processes such as the operation of production facilities, personal transportation or typical action sequences in kitchens and restaurants; thus, conveyor belts, cars and corkscrews also have technical functions.

These considerations indicate how mechanism sketches – as functional analyses – might be developed into more complete representations for the purpose of *controlling* a phenomenon. Often, this control will take the form of not just *representing* but *bringing into being* or tweaking a mechanism that can reliably produce, maintain or prevent a phenomenon. Development of a mechanism sketch can contribute to this, for instance, by specifying an *operational principle* (Vincenti 1990), by identifying clear and reliable *design rules* (Wilholt 2006), or by providing

---

[15] Houkes and Vermaas (2010, Chs. 2 and 4) offer the corresponding definitions of design, use and technical functions.

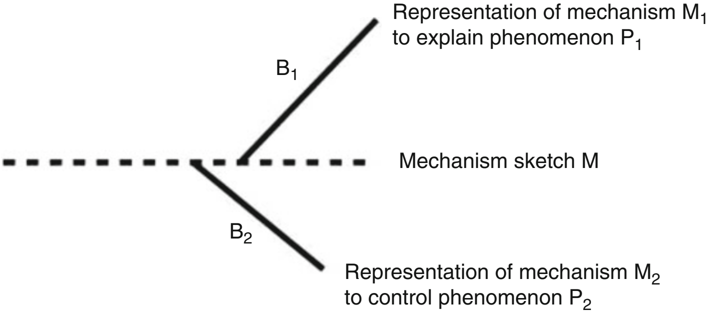


Fig. 12.1 Simple branching of an explanation-oriented research project (*B_1*) and a control-oriented research project (*B_2*), from a shared 'foundational' project

prescriptive knowledge in the form of *technological explanations* (Houkes 2009, Section 8). All such results connect sequences of human actions or interventions to (technologically desirable) outcomes. They add entities and activities to an incomplete mechanism representation, or replace functional placeholders with descriptions of components and specifications of physical processes and tasks for human operators.

The production, maintenance or prevention of these phenomena are, by and large, not a matter of scientific interest alone. Insofar as combustion, artificial vacuums and inoculation are not just explained or predicted but also controlled, they typically serve extra-scientific goals such as societal or commercial interests. In these cases, it is possible that development of mechanism sketches for control of phenomena *combines* the pursuit of 'external' goals with the pursuit of scientific goals – completion of the representation of mechanisms.

A mechanism sketch may thus be developed into different directions, leading to various more complete mechanism representations. Then, each program of development forms one branch among many options. In some branches, development might be aimed exclusively at predicting or explaining phenomena; in others, the aim might include control. Since the latter might develop the sketch differently than the former, control-oriented branches do not need to succeed "completed" explanation- or prediction-oriented branches; in principle, they can be pursued simultaneously by different groups of researchers. In the most simple arrangement (Fig. 12.1), branch B_1 is an explanation-oriented effort; and B_2 is control- or application-oriented. Each employs earlier work on the same mechanism sketch. Further work on this sketch may continue in parallel to work in both branches. In B_1, one would expect a linguistic or diagrammatic representation of the mechanism for purposes of explanation or prediction, whereas the application-orientation of B_2 may be reflected in operational principles of artificial items, design rules or manufacturing procedures. For purposes of illustration, B_1 might represent a branch of research on nuclear fusion that seeks to explain gamma-ray bursts from neutron stars; and B_2 a branch that seeks to control gamma radiation from JET tokamaks. These branches minimally share the use of basic principles of magneto-hydrodynamics, but both need to develop them significantly, and in significantly different ways because of their diverse domains (neutron stars versus JET tokamaks) and, most importantly, basic aims (explanation versus control).

This simple model reflects a sharp division of labour between ‘pure’ and ‘applied’ science, where both are branches of initial work on the mechanism sketch, which one might call ‘basic’ science. This is unlikely to be a plausible representation of interactions between real research efforts. Still, more complicated reconstructions are easily imagined: a branch could, for instance, develop multiple sketches; or work in one branch could stimulate work in another through transfer of mechanism schemas. It would be interesting to test which complications minimally need to be added to model some set of actual projects within fusion research, or other areas that overlap with established images of ‘basic’, ‘pure’ and ‘applied’ science. In the next section, the model will be illustrated and developed in another direction, in order to show how application-oriented research may combine scientific and external goals.

12.3 A Case of Combination: Perovskite Solar Cells

In this section, I develop and illustrate the branch-formation model of the previous section with a case study drawn from research on photovoltaic technology. I show that the distinction between goals internal and external to science need not correspond to a division between ‘pure’ and ‘applied’ branches. Rather, both types of goals can be pursued *within single*, *application-oriented branches*.

In Sect. 12.3.1., I give a short review of developments in photovoltaics, focusing on the recent ‘perovskite’ revolution; in Sect. 12.3.2, I reconstruct this application-oriented research area with a branch-formation model, and bring out how both the formation of new branches and developments within each branch are shaped by aims internal and external to scientific research.

12.3.1 The Road to Perovskites

Cheap, efficient and durable photovoltaic technology is of immediate societal relevance, since it is an important factor in the transition to renewable energy. As all societally valued technologies, a ‘good’ solar cell needs to satisfy a large number of criteria, some of which are likely to be incompatible. Among many other things, a solar cell should convert sunlight into electricity with a high efficiency, it should be durable, yet relatively cheap to produce in large numbers, and it should be possible to produce it with a minimum amount of waste and hazardous materials. The first three of these – efficiency, durability and cost – are often labelled the ‘Golden Triangle’ of photovoltaic technology.

Because of the societal stakes and conflicting specifications involved, photovoltaics are an area of intensive and highly variegated scientific research, at the intersection of solid-state physics, chemistry and engineering. Silicon-based solar cells are a primary focus of attention, since they currently come closest to being competitive with other sources of electricity. Theoretically, single-junction solar cells could have an efficiency of around 34 %, the so-called Shockley-Queisser limit.[16] Commercially available, mass-produced silicon-based cells reach efficiencies of about 22 % under typical (i.e., non-laboratory) conditions of use. Still, there is substantial room for improvement,[17] both in the performance of silicon-based cells over their lifetime and in the production process, which requires several hazardous (e.g., toxic and/or explosive) materials.

Besides improving silicon-based technologies, researchers are investigating a variety of alternatives, often collected under the heading of “emerging photovoltaics”. A major subset of these are dye-sensitised solar cells (DSSCs). These are relatively cheap to produce in comparison with silicon-based cells, but they have been so far less durable and efficient, with conversion rates struggling to reach 10 % under laboratory conditions. Therefore, DSSCs and other emerging photovoltaics used to be more a focus of scientific attention than of industrial or immediate societal interest.

This has changed dramatically in the last few years (Hardin et al. 2012). In 2012, it was reported that cells using materials with so-called ‘perovskite’ crystalline structure are remarkably efficient (Lee et al. 2012). Perovskites, materials with crystalline structure ABX_3 (Fig. 12.2), have all kinds of properties, such as ferroelectricity and super- and semi-conductivity, that make them of interest to a broad variety of purposes. Using them in photovoltaic cells was not only relatively cheap and easy. It also led, in a few steps discussed later, to spectacular increases in efficiency: from only 3.5 % in the first experimental attempts from 2009, to 10 % in 2012, almost 15 % in 2014 – and over the magical barrier of 20 % early in 2015.[18] In a field where single-decimal increments are worth reporting, this is revolutionary progress. Perovskite solar cells hold a real promise of being commercially viable, offering similar efficiency as silicon-based cells at a much lower cost of materials. Interest among researchers in the field and the general public alike has boomed. Lee et al. (2012), the *Science* paper that announced the revolution, has already been cited over 1300 times.[19] In a 2014 IEEE Spectrum item with the telling title “Perovskite is the New Black in the Solar World”,[20] Henry Snaith – a perovskite pioneer – claims that every photovoltaics research group in the world is currently looking at perovskite

---

[16] There are ways to exceed the Shockley-Queisser limit. One is to use ‘tandem’ cells rather than single-junction ones. These combine junctions that are sensitive to different parts of the solar spectrum, and can reach efficiencies of over 50 % in unconcentrated sunlight.

[17] Making explicit how much room for improvement there is requires developing appropriate life-cycle assessments of environmental impact, which are themselves a major area of interest and controversy.

[18] As reported at http://spectrum.ieee.org/energywise/green-tech/solar/perovskite-solar-cell-bests-bugbears-reaches-record-efficiency (item dated January 2015; accessed August 2015)

[19] Google Scholar (August 2015).

[20] http://spectrum.ieee.org/green-tech/solar/perovskite-is-the-new-black-in-the-solar-world (accessed August 2015).

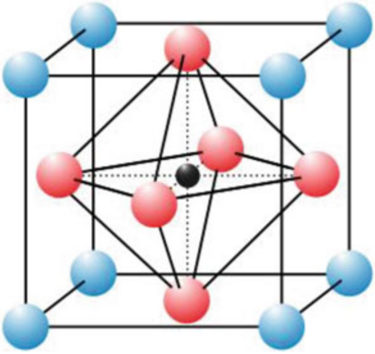


Fig. 12.2 Perovskite crystalline structure of a material ABX_3 composed of cations A and B of different sizes (*blue and black) and anions X binding to both (red) (Source: Wikimedia Commons)*

cells, and that commercial perovskite cells will be available in 2017. The major obstacles to be cleared are durability, since perovskite materials are as yet too sensitive to moisture and prolonged exposure to sunlight to last the required 25 years; and the use of hazardous materials, such as the lead used as A cation in many perovskite cells.

To understand which steps have been taken requires some information about how photovoltaic cells work. Most generally, these cells require three activities: photo-excitation/electron donation; electron acceptance; and transportation of electrons and holes.[21] In terms of entities, a first component of the system enters an excited state by light absorption, and transfers electrons to a second component via a third, intermediate component. In silicon-based cells, two materials suffice. A layer of n-doped silicon is the electron donor. It is immediately adjacent to a layer of p-doped silicon, which is the acceptor; electrons and holes are transported by both materials. By contrast, a dye-sensitized solar cell is unique in using a separate component and material for each of the three central activities (Hagfeldt et al. 2010, Section 3).[22] Light is absorbed in a dye that coats a metal oxide. The electrons generated by this absorption are transferred from the dye to an electrolyte fluid,[23] which fills pores in the dyed metal oxide and which regenerates the dye (i.e., brings it back to its initial state) before it can recapture the electron. Finally, the electrons are transported by diffusion to the metal oxide, which acts as the electron acceptor (Fig. 12.3).

Having three separate materials and components might be a successful, modularity-based way of searching for the optimal design of a photovoltaic cell. However, it also introduces an additional interface where items might interfere with each other's performance, as well as additional possibilities for 'weak links' and

---

[21] These are the 'inner' workings only. Effective photovoltaic cells also require conduction to an external load; protection of sensitive materials combined with transparency to sunlight; etc.

[22] In fact, making explicit the transportation role – which is trivially played by the n-type and p-type layers in silicon cells – is pivotal to understanding how DSSCs work. This shows that, even though photovoltaic cells may be represented by the same mechanism sketch, this sketch may need to be re-arranged in order to be developed in some directions.

[23] Some dyes not only play the electron-donor role, but also that of electron acceptor, facilitating transfer to the metal oxide.

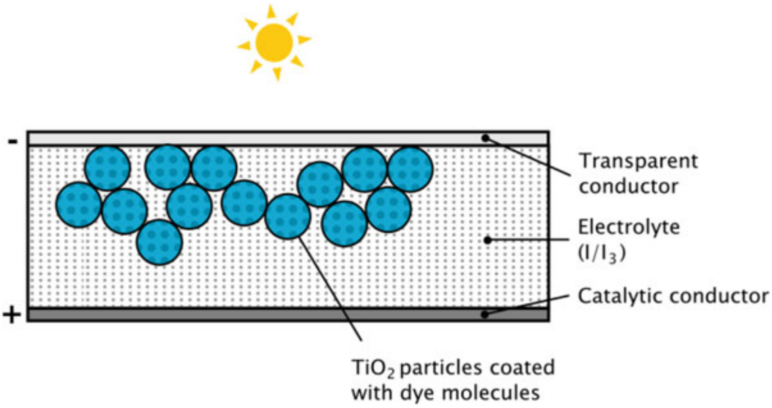


Fig. 12.3 Diagram of the main components of a dye-sensitised solar cell (Source: Wikimedia Commons)

loss of overall performance. The liquid electrolyte, for instance, may freeze at low temperatures, potentially damaging the cell. Researchers have tried different metal oxides (besides the typical titanium oxide), electrolytes (besides the typical iodide) and especially dyes, in different arrangements and with a variety of additional functional items, e.g., to prevent recapture of electrons by the dye; see (Hagfeldt et al. 2010) for a review of such attempts.

Around a decade ago, perovskite materials were first used as electron donors. This attempt was driven by both cost and performance issues with other dyes. Until recently, many of the most efficient DSSCs used dyes containing ruthenium, which is relatively rare[24] and therefore expensive. Organic dyes can be used as replacement, but their narrow absorption bandwidths lead to far lower efficiencies. Some perovskites, such as methylammonium lead halides ($CH_3NH_3PbI_3$), have a large bandwidth while still being cheap to produce, making them natural replacements for ruthenium-based and organic dyes. Unfortunately, these materials tend to decompose in the electrolytes within minutes – failing the third, durability criterion in the Golden Triangle of photovoltaics. This was addressed by using solid-state materials as replacements for the electrolytes, which are hazardous during production and temperature-sensitive during operation anyway. The result was, in many ways, a conventional DSSC, albeit with two non-standard functional items – the perovskite absorber and the solid-state transporter – and an exceptional efficiency of some 10 % (Lee et al. 2012).

This application of perovskites in a DSSC led to a next stage, when it was found that there was electron transport *within* the perovskite absorber. For application in a DSSC, this is an undesirable property, because it reduces transfer to the metal oxide, the designated electron acceptor, and therefore reduces the efficiency. However,

---

[24] Rarity is not just problematic because it leads to higher production costs, but also because it limits how much power could be generated by photovoltaic cells of this type.

researchers realized that, because of this property, perovskites could have *all three* basic functions of items in a photovoltaic cell. As Snaith writes, in these perovskite cells, “there no longer remain any of the original components of the DSSC; therefore, we can consider this an ‘evolutionary branching point’” (2013, p. 3625).[25] Most importantly, from a practical perspective, is that perovskites appear to have come into their own in this one-material show: the efficiencies of 15 % and 20 %, reported in 2014 and 2015 respectively, were achieved with all-perovskite cells. Therefore, these cells have not only left the DSSC architecture behind, but also far outstripped their performance. Still, durability and, to a lesser extent, the toxicity of some materials (in particular the lead used in many perovskites) remain major obstacles.

12.3.2 A Branch-Formation Reconstruction of Photovoltaic Research

The steps that led to the development of perovskite solar cells can be represented as the formation of successive branches from a shared mechanism sketch. This sketch indicates the three basic activities and leaves open a host of details: which items perform these activities; which other activities might be needed; which properties of the items and environmental factors might interfere; even the basic mechanism of electron and hole transportation.

The sketch can be straightforwardly understood as a functional analysis, in which items are ascribed the functions of ‘photoexcitation’ (f_{PE}), ‘electron acceptance’ (f_{EA}) and ‘transportation’ (f_{TR}), without describing their physical and chemical structure or their geometrical organization (Fig. 12.4).[26] For the various types of photovoltaic technology, this functional analysis is replaced with a more complete representation of the mechanism. This includes more details about the entities involved, such as the physical and chemical properties of constituent materials and the ‘normal’ configuration of devices; and it also develops the representation of activities, e.g., by distinguishing various intermediate steps and by describing the kinetics and energetics of various processes involved. The representation might also include new function ascriptions, for instance, by specifying a ‘scaffolding’ for the perovskite material that has all three basic functions.

Because they involve different developments of the same functional analysis, research on different types of photovoltaic technology can be naturally represented as branch formation. Development of silicon-based cells can be modelled as a

---

[25] Metal oxides are still used, but as a scaffolding for the perovskite rather than electron acceptors.

[26] In line with the ICE-theory of functions (Houkes and Vermaas 2010, Ch. 4), the three central activities in photovoltaic cells are taken as corresponding to functions ascriptions to items, since they are the result of deliberate design, rather than functional roles, which feature in post-hoc explanations of systemic behaviour.

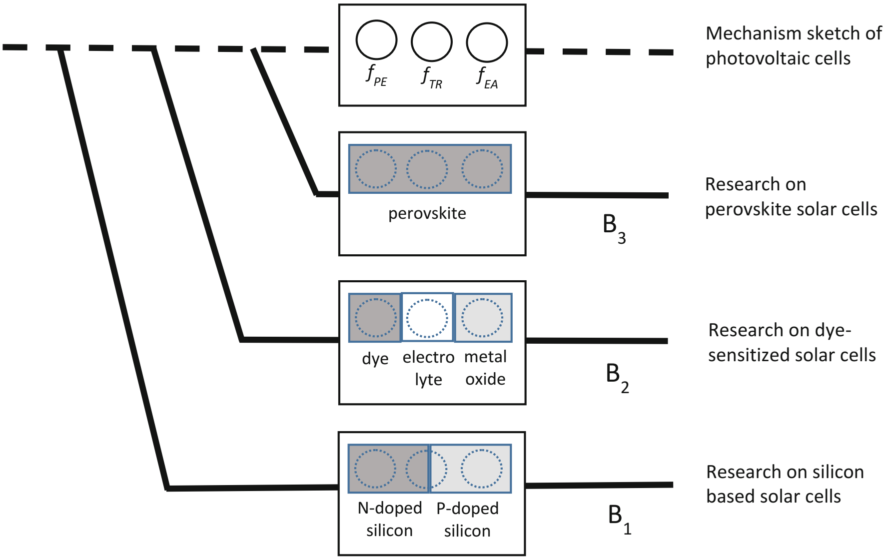


Fig. 12.4 A branch-formation model of research on photovoltaic technologies. In each branch, research aims at explanation and control of a mechanism that consists of different materials (*greyscale rectangles*) that are ascribed the three basic functions of photovoltaic devices (*dotted circles*). Representations of these mechanisms develop a sketch that specifies these basic functions (*solid circles*)

project B_1 branching out from the mechanism sketch.[27] Here, the functions f_{PE} and f_{EA} are ascribed to n-doped and p-doped silicon respectively and the function f_{TR} to both. Work on dye-sensitised solar cells forms another, more recent branch B_2, in which dyes (ruthenium-based, organic or perovskite) have the photoexcitation function, metal oxide that of acceptance, and the transporters are electrolytes. Finally, perovskite cells are the focus of interest in the most recent branch B_3, in which perovskite materials are used for photoexcitation, acceptance and transportation.

The more detailed representations of these mechanisms in photovoltaic cells have the central characteristics of mechanistic explanations, such as the identification of component parts and activities, and a representation – often diagrammatic rather than linguistic – of their spatial and temporal organization (Bechtel and Abrahamsen 2005). These explanations are also a means for enhancement, however: researchers seek to understand the properties of and interactions between components, and the characteristics of processes, in order to *improve* the performance of devices – in particular their efficiency and durability. The point is not just

---

[27] As indicated above, *all* photovoltaic cells were, until Graetzel's pioneering work on DSSCs, understood on the basis of silicon-based cells, without explicit representation of the transportation function. Thus, it would be historically more accurate to represent the tripartite functional analysis as a mechanism *schema*, constructed from a representation of the mechanism of silicon-based cells.

to understand what goes on in a photovoltaic cell, but also why it is a poor or promising cell, and to what extent it can be improved.

To give a general example of such ‘explanations-to-enhance’, representations of the energetics of various processes in energy-level diagrams provide explanations of overall properties of photovoltaic cells. Yet this explanation serves as a means to identify bottlenecks in the conversion efficiency and possibilities for improvement. A more specific example is provided by Snaith, when he writes about one of the central aims of research on perovskite cells:

> the real challenge to make sure that this technology has a dramatic impact on the solar industry is to focus on continually understanding and enhancing the stability. (…) [U]nderstanding changes that can occur during aging and developing components and processing steps that enhance stability should be one of the key scientific goals for the field. (2013, p. 3628)

What is telling is, first, that understanding/explaining and enhancing/developing are identified as joint *scientific* goals; and, second, that they are identified as key contributions to a *societal* challenge, namely revolutionizing the solar industry. External goals therefore play an important guiding role in photovoltaic research: not only efficiency – a notion in which societal values might play a role, but often implicitly; and which does come in almost purely physical forms (Alexander 2009) – but also the costs and methods of production and the toxicity of materials – which are clearly related to the context of application and societal role envisaged for photovoltaics – shape developments of branches and within branches. External goals and non-epistemic values have been *internalized* in photovoltaic research, as was emphasized by the finalization programme.

In contrast with the tenets of that programme, this internalization does not succeed or build upon closed theories. In photovoltaics, there are no “laws from which sufficiently precise and reliable predictions can be derived” (Böhme et al. 1976, p. 309); or, in more mechanistic terms, there are no sufficiently complete representations of mechanisms for the purposes of either explanation or control prior to photovoltaic research. Rather, developing such representations is central to research on (a particular type of) photovoltaic cells. Explaining in detail how perovskite materials transport electrons and holes, constructing an accurate account of the kinetics and energetics of a perovskite cell, and understanding how such a cell may be affected by environmental conditions are central goals. Most publications in the field report contributions to these goals, and are difficult to distinguish in this respect from publications in other areas of physical chemistry and solid-state physics. Thus, photovoltaic research does, and needs to, involve pursuit of traditional scientific goals – explanation; modelling; understanding – albeit not in isolation from pursuit of external goals. It is application-oriented research rather than applied (i.e., theory-applying) research.

Because mechanistic understanding of photovoltaic cells is incomplete, and made more complete, in every branch, there is another way in which the combination of internal and external goals can be brought out. All branches are guided by the same three overriding external goals or non-epistemic values – the Golden

Triangle of photovoltaic technology. Likewise, all branches share a mechanism sketch – the basic, tripartite functional architecture of a photovoltaic cell. This means that, in this particular field, branches are relatively straightforwardly recombined. Whereas combinations of different mechanism sketches require overarching sketches or other ways of understanding interfield relations,[28] and combinations of different external goals require reconsideration of every trade-off made earlier, branches that share both external and internal goals merge naturally, and the drivers of such mergers are likely to be the external goals. Indeed, hybrid multi-junction solar cells that combine single-junction cells of different types are one way of exceeding the Shockley-Queisser limit; consequently, hybrid "tandems" of silicon-based and perovskite cells are considered as a means for reducing the cost of existing silicon-based cells without impeding efficiency.[29] The diagram in Fig. 12.4 only brings out how research efforts branch out from a shared sketch; a more complete reconstruction would also include reticulations.

12.4 Conclusions and Outlook

In this paper, I have presented a model of application-oriented science. This 'branch-formation' model modifies the central ideas of the Starnberg finalization programme using elements from more recent, mechanistic philosophy of science and the philosophy of technical artefacts. It represents research efforts as branches that develop more complete representations of mechanisms, and that may share rudimentary versions of these representations – mechanism sketches – with other branches. Developments of branches and within branches may be guided by traditional scientific goals such as explaining and predicting phenomena; by external, societal goals such as safety and sustainability through controlling phenomena in artificial devices; or by both types of goals. As an illustration of the latter, dual-purpose research, I reconstructed the development of perovskite solar cells as a recent branch in photovoltaic research. This area can be characterized with a mechanism sketch (a tripartite functional analysis) shared by several branches. Each branch seeks to complete the representation of the operating mechanism of a particular type of photovoltaic cell, in order to enhance the performance and durability of this cell, while minimizing the cost of production.

The branch-formation model allows a detailed reconstruction of knowledge accumulation in research areas, by focusing on the completion of mechanism representations. Furthermore, it brings out how external goals may affect this accumulation,

---

[28] Interfield relations are a prominent topic of inquiry in mechanicist philosophy of science; see, e.g., Craver and Darden (2013, Chap. 10) for an introduction.

[29] The need to reformulate the shared functional architecture with the discovery of DSSCs and recurrent discussions of the most appropriate way of operationalizing performance show that, even in photovoltaic research, internal and external goals of different branches are only approximately identical.

by reconstructing application-oriented branches; and it shows how external goals may be pursued without attenuating internal goals: in some cases, development of mechanism representations may serve both control and understanding. As such, the model offers an epistemological micro-perspective on scientific research in which efforts can, in principle, be evaluated positively with regard to both internal and external goals.

This supplements existing influential approaches to application-oriented research in science-policy studies, which typically focus on entire research areas rather than specific efforts, and offer little support for an evaluative perspective. Both the NPK and 3H approaches would focus on photovoltaic research in general. For this, they would rightfully emphasize its transdisciplinary character, combining physics, chemistry and engineering science. Moreover, the 3H approach in particular could target its organizational form and the underlying incentives for researchers. This contrasts with and supplements the focus on the branch-formation model, which necessarily ignores – for instance – the close connections between research on and commercialisation of photovoltaic cells: that Snaith has not only done pioneering research on perovskite cells, but is also co-founder of Oxford Photovoltaics, which seeks to commercialise such cells, cannot be represented in terms of completion of mechanism sketches. Conversely, possible conflicts between commercial interests and knowledge accumulation can be analysed by studying the particular industry-university-government relationship involved.

It should be stressed that the branch-formation model is a *model*, meant to bring out epistemic aspects of application-oriented science. In the form presented here, the model focuses on the development of new branches, and on developments within branches; it leaves room for, but does not reconstruct, recombination of branches. Furthermore, it has only been illustrated with one case study, as a first ‘indication of concept’ rather than as evidence for its accuracy. Applications to other cases are likely to lead to further development and modification.

Most importantly, the branch-formation model mainly plays the first supplementary role identified in Sect. 12.1: it provides a micro-perspective on knowledge production in particular areas of scientific research, and it aids evaluation in identifying the relative importance of internal and external goals as well as the possibility of combining them in a single area. As such, it provides materials for developing the second type of supplement – an evaluation of overall changes in scientific research – but it does not act as such a supplement. To show this limitation, as well as to bring out the usefulness of the model in framing an overall evaluation, I will close by outlining three scenarios for the re-orientation of scientific research that are very different in their evaluative implications, and are all compatible with the branch-formation model.

A first, optimistic way to explicate the intuition that scientific research is *increasingly* focused on producing knowledge or other results that are directly relevant to non-scientists is to highlight combinatory branches. If more and more branches in the science system would show a duality of purpose rather than a focus on explanation and prediction alone, a re-orientation towards applications is in principle possible without diminishing the production of scientifically relevant knowledge.

Pursuit of scientific goals such as explanation and prediction might proceed unhindered, even if scientists increasingly *also* pursue external goals: re-orientation amounts to a broadening of goals rather than a switch from one set of goals to another. An alternative, more pessimistic construal of the re-orientation is that promising branches that primarily or exclusively serve goals internal to science fail to develop, in favour of branches that are guided by both internal and external goals, or even exclusively by external goals. Then, the overall structure of the 'tree' or set of branches would change as a result of changes in incentives; the tree would – metaphorically speaking – move along with a source of light. A third, still different and equally pessimistic scenario is that an orientation on application might not detract from the pursuit of internal goals in completing mechanism representations, but that it might put a premium on working in the later stages of developing representations rather than the earlier. Then, the shift is not from 'scientific' to 'mixed' or 'applied' branches, but rather away from potentially ground-breaking work on mechanism sketches. In this case, changes in lighting might not just affect the overall shape of the tree, but also affect its long-term growth perspectives.

The branch-formation model plays a role in distinguishing these scenarios; and the case study shows that the most pessimistic epistemological perspective on the application re-orientation – that pursuit of external goals can only diminish the scientific value of research activities – is unwarranted. Yet the other scenarios, which are both compatible with the model and the case study, show that there is still ample room for pessimism. This essay only takes a small first step in an evaluation of the ongoing changes in the science system, and it indicates some of the difficulties in making an overall judgement. Work on further case studies, and perhaps more sophisticated versions of this model, is required to find out to which – if any –of the three scenarios identified here is the most accurate.

References

Ahrweiler, P. (1996). Negotiating a new science. In H. Etzkowitz & L. Leydesdorff (Eds.), *Universities and the global knowledge economy* (pp. 97–105). Amsterdam: Thomson Learning.

Alexander, J.K. (2009). The concept of efficiency. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 1007–1030). Amsterdam: North Holland.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanistic alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421–441.

Biddle, J. (2012). State of the field: Transient underdetermination and values in science. *Studies in History and Philosophy of Science, 44*, 124–133.

Böhme, G., van den Daele, W., & Krohn, W. (1976). Finalization of science. *Social Science Information, 15*, 307–330.

Böhme, G., van den Daele, W., Hohlfeld, R., Krohn, W., & Schäfer, W. (1983). *Finalization in science*. Dordrecht: Reidel.

Craver, C. F. (2006). When mechanistic models explain. *Synthese, 153*, 355–376.

Craver, C. F., & Darden, L. (2013). *In search of mechanisms*. Chicago: Chicago University Press.

Daston, L., & Galison, P. (2007). *Objectivity*. Brooklyn: Zone Books.

Douglas, H. E. (2009). *Science, policy, and the value-free ideal*. Pittsburgh: University of Pittsburgh Press.

Etzkowitz, H. (2002). *MIT and the rise of entrepreneurial science*. London: Routledge.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation. *Research Policy, 29*, 109–123.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge*. London: Sage.

Godin, B. (1998). Writing performative history: The new *New Atlantis*? *Social Studies of Science, 28*, 465–483.

Hagfeldt, A., Boschloo, G., Sun. L., Kloo, L., & Pettersson, H. (2010). Dye-sensitised solar cells. *Chemical Review*, 110, 6595–6663.

Hardin, B. E., Snaith, H. J., & McGehee, M. D. (2012). The renaissance of dye-sensitized solar cells. *Nature Photonics, 6*, 162–169.

Hessels, L. K., & van Lente, H. (2008). Re-thinking new knowledge production. *Research Policy, 37*, 740–760.

Houkes, W. (2009). The nature of technological knowledge. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 309–350). Amsterdam: North Holland.

Houkes, W., & Vermaas, P. E. (2010). *Technical functions: The design and use of artefacts*. Dordrecht: Springer.

Irzik, G. (2010). Why should philosophers of science pay attention to the commercialization of academic science? In M. Suárez, M. Dorato, & M. Rédei (Eds.), *EPSA epistemology and methodology of science* (pp. 129–138). Dordrecht: Springer.

Kitcher, P. (2001). *Science, truth and democracy*. New York: Oxford University Press.

Lee, M. M., Teuscher, J., Miyasaka, T., Murakami, T. N., & Snaith, H. J. (2012). Efficient hybrid solar cells based on meso-superstructured organometal halide perovskites. *Science, 338*, 643–647.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

Mirowski, P. (2004). The scientific dimensions of social knowledge and their distant echoes in 20th century American philosophy of science. *Studies in History and Philosophy of Science, 35*, 283–326.

Mirowski, P., & Sent, E.-M. (2008). The commercialization of science and the response of STS. In E. J. Hackett, O. Amsterdamska, M. Lynch, & J. Wajcman (Eds.), *The handbook of science and technology studies* (3 revth ed., pp. 635–689). Cambridge, MA: The MIT Press.

Piccinini, G., & Craver, C. F. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese, 183*, 283–311.

Radder, H. (Ed.). (2010). *The commodification of academic research*. Pittsburgh: University of Pittsburgh Press.

Shinn, T. (2002). The triple helix and new production of knowledge: Prepackaged thinking on science and technology. *Social Studies of Science, 32*, 599–614.

Snaith, H. J. (2013). Perovskites: The emergence of a new era for low-cost, high-efficiency solar cells. *The Journal of Physical Chemistry Letters, 4*, 3623–3630.

Vincenti, W. G. (1990). *What engineers know and how they know it*. Baltimore: Johns Hopkins UP.

Wilholt, T. (2006). Design rules. *Philosophy of Science, 73*, 66–89.

Ziman, J. (2000). *Real science*. Cambridge: Cambridge University Press.

Wybo Houkes holds an MSc in Physics from the Free University of Amsterdam, and an MA and PhD in philosophy from the University of Leiden. He is associate professor of philosophy of science and technology at Eindhoven University of Technology. Together with Pieter Vermaas, he authored *Technical Functions* (Springer, 2010). He is a former NWO-Vidi Fellow for the project "Darwinism in the Human-Made World" (2008–2013). His main current topics of interest include: the nature of technological knowledge; the growing importance of societal and practical relevance to scientific research; the use of biological metaphors and techniques in understanding and designing technology; the ethics and metaphysics of artefact copying. Some current projects: *Domesticating Darwinism* (in progress); the impact of digital technology on creative work practices in architectural firms; the robustness of model-based demographic explanations of cultural and technological change.

Chapter 13
Methodological Classification of Innovative Engineering Projects

Sjoerd D. Zwart and Marc J. de Vries

Abstract In this chapter we report on and discuss our empirical classification of innovative engineering projects. Basic innovative engineering projects are characterized by their overall goal and accompanying method. On the basis of this goal and method, we classify engineering projects as all falling in one of the following categories: (1) *Descriptive knowledge* as prevalent in the descriptive sciences; (2) *Design* of artefacts and processes; (3) *Engineering Means-end knowledge*; (4) *Modeling* (simulation serious gaming included); (5) *Engineering optimization*; and (6) *Engineering mathematics*. These categories are illustrated with examples drawn from our educational experiences. Formally our classification system is a partition: the categories are mutually exclusive and collectively exhaustive. Regarding its empirical power, we *claim* intra-departmental completeness for the projects that we have studied at the Departments of Mechanics and Applied Physics of Delft University of Technology; we *hypothesize* intra-academic completeness within Universities of Technology; and we hope for and *encourage* investigating extra-academic completeness regarding engineering in industry. Besides having significant consequences for the methodology of the engineering sciences, our categorization provides a new way to study empirically the relation between science and technology.

Keywords Types of engineering projects • Engineering methodology • Descriptive knowledge • Design • Engineering means-end knowledge • Modeling • Optimization

---

S.D. Zwart (✉) • M.J. de Vries
Delft University of Technology, Delft, The Netherlands

Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: s.d.zwart@tudelft.nl; m.j.devries@tudelft.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_13

13.1 Introduction

A core question in the philosophy of technology is the distinction between science and technology, and more in particular, between engineering and performing scientific activities (Meijers 2009). One possible way of approaching this issue is by looking at the kinds of knowledge produced in both fields. Important contributions from this perspective have been made by Layton in his seminal paper *Technology as Knowledge* (1974), and by Walter Vincenti in his ground breaking (1990) book, *What engineers know and how they know it*. Despite the great expertise that characterizes the work of Layton, Vincenti and their followers, the results of their work do not settle the question about the relation between engineering and science and the different ways of knowing occurring in either (Houkes 2009).

Another perspective takes engineering to be technological problem solving. This perspective, which has its defenders in philosophy and didactics (Koen 2003; Sheppard et al. 2007, p. 431; Graaff and Kolmos 2003), also has a respectable tradition within the engineering world (Krick 1969; Hubka and Eder 1996, Chap. 1; Pahl et al. 2007, Sect. 2.2). One of the underlying ideas of this view concerns the difference between the aims of science and technology: the final goal of the former is descriptive knowledge and of the latter well-functioning technical artefacts or processes. This conceptual distinction has a rich history that goes back at least to the contrast between *epistêmê* (scientific knowledge) and *technê* (craft or art) as used for instance by Aristotle in his *Nicomachean Ethics* (Book VI). Despite its plausibility at first sight, the problem-solving perspective has been heavily criticized, although with sparse empirical evidence from engineering practices. This leaves us with the impasse of acknowledging two concepts, science and technology, with major accompanying societal institutions, but without being able to characterize their differences.

In the programmatic introduction to their (2000), Kroes and Meijers call for an empirical turn within the philosophy of technology, a discipline, which at the time (and even perhaps today) they considered in need of a unifying direction. In this introduction, the authors argue for a philosophy of technology that is analytic and conceptual, empirically well informed, and possibly normative. Philosophers should open the black box of technology and take into account the detailed differences within the various technological fields. The main underlying reason for this advice is the strong non-monotonicity of normative judgements. Social criticism of technology, which had been in vogue in philosophy of technology for a long time, was mainly based on an external view of technology. Addition of information from the contents of the black box could readily render the external assessments implausible and obsolete.

Answering Kroes and Meijers' call for an empirical turn, we aspire to tackle the impasse mentioned above by opening the black box and investigating the projects carried out within scientific and engineering practices. We take these projects to be defined by the effective goal and the accompanying methods as employed in real-world practice. In this chapter we start with investigating the engineering side of the problem, leaving the analyses of scientific projects to another occasion. It may turn out that the types of projects within science and technology will overlap. Nevertheless, we expect the distribution of the types of projects within the two realms to differ considerably. As an additional advantage of the empirical turn, Kroes and Meijers mention that engagement with 'new topics and new conceptual frameworks […] will give rise to their own, specific kinds of philosophical issues' (2000, p. xxv). As we will see, in this chapter this 'prophecy' will also be fulfilled.

Overall, in this chapter our research objective is to present an empirically adequate and theoretically sound classification of innovative engineering projects that is based on the *type of goal* and the *method* used to achieve that goal. We call a classification empirically adequate if it covers all the project types encountered and 'cuts the world at its joints', covering all these types in a natural way. We conceive engineering projects as technological problem-solving activities, which are considered innovative if they have not been carried out before. That is, no standardized recipes exist that, when being followed, will achieve the goal of the project 'semi-automatically'.

Our research method is mainly empirical and based on engineering practices within academia. We have classified hundreds of Bachelor End Projects (abbreviated: BEPs) carried out at the Delft University of Technology. Our method also has theoretical ingredients, as we use standard insights from the methodological literature. Our investigations have resulted in a classification that, as we hypothesize, is intra-academically complete, i.e., it categorizes all academic innovative engineering projects. It remains to be investigated, however, whether it is complete regarding engineering projects undertaken in industry.[1] We show that besides descriptive knowledge, modeling and mathematics – which are all important for both science and engineering – design, engineering means-end knowledge and optimizing are other major aims of engineering projects. They may turn out to be goals of lesser importance in the sciences. Besides being relevant for the philosophy of technology, our choice to start with academic projects has the advantage of helping students and engineering professionals to formulate clearly the purpose of their projects and to be explicit about their methods.

Finally, we wish to make two terminological remarks. First, in this chapter the terms 'categorization', 'classification (system)', 'taxonomy' and 'typology' will be treated as synonyms. Although in the literature a distinction is made between taxonomies as being empirical and typologies as being conceptual, we ignore the difference. However, we do distinguish between these four concepts on the one hand, and on the other a *partition*, being a set of mutually exclusive and collectively exhaustive subsets.

Second, methodology as the comparative study of methods has traditionally been closely connected to methods used to obtain descriptive scientific knowledge

---

[1] Our choice for starting with academic projects has the following advantage regarding extrapolation: if the extra-academic engineering practices fit our classification well, we have a convincing validation; if they do not fit our classification we have shown that our engineering education system is minimally incomplete.

(Creswell 2013; Wilson 1952). Here, the phrase ‘methodology of a project’ refers to the project strategy applied to achieve some predefined but not necessarily cognitive goal. The engineering literature encompasses many *methods and techniques*, which are applied while carrying out specific (engineering) tasks within specified contexts (QFD, morphological analysis, MCDA, Pugh Charts etc.). In this chapter we extend the connotation of the term ‘methodology’ from the study of methods applied in the descriptive sciences toward the general study of all methods and techniques used to solve problems in engineering and science. Consequently, the body of engineering methods also comprises many ways in which engineers realize designs of artifacts and processes, which are often referred to as ‘design methods’. These methods belong therefore to what we identify as *engineering methodology*. Thus, we generalize the more traditional notion of methodology, associated mainly with methods to obtain descriptive knowledge, into the engineering notion of methodology, which also comprises methods geared to other forms of engineering problem solving.

In the next section we present a preliminary, more technical description of our classification. In the Sects. 13.3, 13.4, 13.5, 13.6, 13.7, and 13.8, we explain the individual categories in depth: descriptive knowledge, design, engineering means-end knowledge, modeling, engineering optimization, and engineering mathematics. In the penultimate section, we discuss how projects can be combined in larger ones. We end the chapter with a discussion of our results and some indications of possible future research.

13.2 The Classification, a First Encounter

When we rationally reconstructed the historical development of the six categories that we claim serve to classify innovative engineering projects, we realized that, systematically, we have based these categories on the interplay between the main goal of an innovative engineering project and its *accompanying method*. Our criterion is that it must be possible to consider this goal as an *end-in-itself*, and that the overall method should be sufficient to reach that goal and should not additionally serve other purposes. A category is therefore determined by the coherent unity of a type of end-in-itself and its accompanying type of method. We define six types of such pairs of goal and method. For example, *descriptive knowledge* is a type of ultimate goal and the empirical cycle is the undivided accompanying type of method that will achieve this goal. Design is another type of goal, and the design cycle is the type of its accompanying method.[2] If an engineering project has a single categorical goal and no subgoals, and uses the categorical method that corresponds to this goal, we call it an *atomic engineering project* or activity. In practice only a minority of innovative engineering projects are atomic. Most projects are complex and have more than one subgoal.

---

[2] Of course we are well aware that these types of methods are again reconstructions and not actual patterns in daily practices.

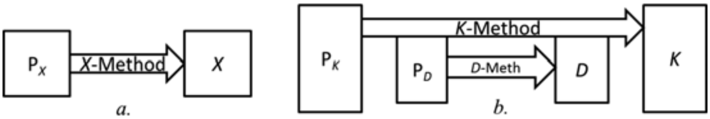


Fig. 13.1 (**a**) The diagrams of atomic projects; (**b**) a molecular project. (*K* knowledge, *D* design). X may instantiate K, D, P, M, F, O abbreviating descriptive Knowledge, Design, Practical means-end knowledge, Modeling, Formal mathematical results and Optimizations, respectively

Means-end relations are often hierarchical. Many innovative engineering projects have one overall goal, and accordingly use one overall method, but they accomplish this ultimate goal by realizing various subgoals, each with its accompanying method. We will call such projects single-purpose or *molecular* projects and we will categorize them by means of their overall goal and method. These projects are built up out of atomic engineering activities. Our classification system thus serves two purposes. It directly categorizes all atomic and molecular engineering projects, and it defines the atomic building blocks of the latter.

Figure 13.1 illustrates the relation between atomic and molecular projects. The *X* covers the six categories of projects introduced above and explained below and the diagram on the right represents an overall descriptive knowledge project that contributes to a larger project in which an artefact (e.g., a measurement instrument) is designed.

Not all engineering projects are either atomic or molecular. Many larger projects are so all-compassing that they have several ultimate goals and use different methods in parallel. On the other hand, many smaller ones have goals that are not ends-in-itself and fail to have a distinctive accompanying method. Thus, some larger projects carry out more than one atomic engineering project in parallel, while other smaller ones achieve only part of an atomic project. Let us have a closer look at both of them.

Larger overarching projects, like for instance the development of a new revolutionary sewage treatment technology, often have different kinds of main goals simultaneously, such as scientific, design and engineering-knowledge goals (Zwart and Kroes 2015). These goals can all be considered ends-in-themselves. We call these parallel enterprises *multiple-purposed* projects. Because of their multiplicity of aims and methods they do not specify or belong to one coherent category. We will characterize multiple-purposed projects using *Radar Charts* (see Sect. 13.9) showing how the time spent is distributed over the various ultimate goals.

On the other side of the spectrum, smaller projects are often characterized by goals that can hardly be considered ends-in-themselves, and are not accompanied by a complete distinctive methodology. Because they are defined by goals that are not ends-in-themselves, these projects do not form atomic projects. We call them *subatomic* projects and we classify them according to the atomic project they serve. For instance, although having a specific goal, a *proof of concept* (also called proof of principle) uses only part of a standard method. Normally, it is an empirical proof of the functionality of some operational principle. Now, if the ultimate goal of a

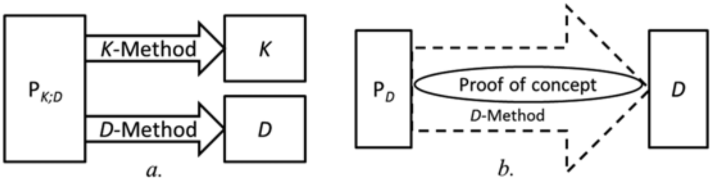


Fig. 13.2 Diagram (**a**) is a multiple purpose design and knowledge project; (**b**) a subatomic project, the proof of concept serves a design

proof of concept is a design, it is considered to complete only half of the design cycle. If, in contrast, its ultimate goal is the testing and underpinning of some engineering-design knowledge, it is considered as falling under engineering means-end knowledge methodology. Consequently, a proof of concept is placed in the category in which the overall project to which it belongs is placed. Being without a specific method, it fails to define a distinct category (Fig. 13.2).

We claim that our six categories cover all innovative engineering projects, at least the sorts of projects undertaken in the departments we have investigated. More precisely, all atomic and molecular projects belong to one of our six categories and all subatomic projects link naturally to one category. Moreover, all multiple-purpose innovative engineering projects consist only of atomic and molecular engineering subprojects falling in one of our six categories.

Let us turn to the characterization of our six categories. The main goal of innovative engineering projects within the first category is to produce trustworthy *descriptive knowledge* (1), which might be considered as possibly true or false. This is the traditional methodology category with the same kind of goal pursued in the natural and social sciences. Its accompanying methodology has the form of the empirical cycle. It consists in finding the relevant dependent and independent variables used to formulate the hypothesis, which, being the provisional answer to the research question, should subsequently be tested.

A completely different goal of engineering projects is a *design of an artifact or process* (2), which is largely absent as ultimate goal – apart from the design of experiments – within the natural sciences. The end-product of such a project is not claimed to be true or false. It is to a larger or lesser extent effective, functional or fulfilling the design requirements. Accompanying methodologies can be reconstructed as a design cycle based on a design brief that states requirements, a working principle, and a morphological overview in which the different design options are systematically compared with respect to the extent to which they meet the requirements.

Projects of the third category aspire to produce (*engineering*) *means-end knowledge* (3). This knowledge explicates how to act in order to achieve some prespecified technological goal in some given context. Class (3) has largely escaped the attention of major methodologists so far. We conjecture that technological means-end knowledge has to be underpinned top-down by way of scientific theories and bottom-up by explicating underlying causal structures or mechanisms.

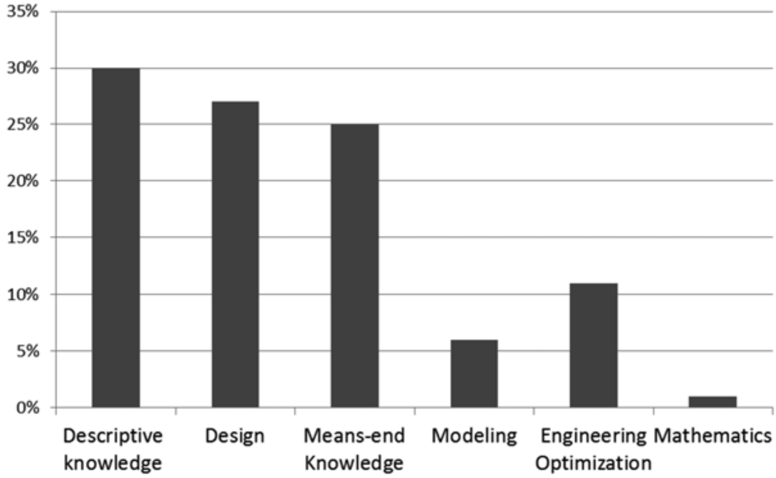


Fig. 13.3 The distribution of the projects over the six categories

Most of the projects we have examined are, roughly, divided equally over the project categories (1), (2) and (3) (see the bar chart in Fig. 13.3). These project categories, however, do not cover the methodological needs of all of them. Among the minority of projects left, we identified three other ones. Some of them are clearly *modeling* projects (4). They aim at the production of a model and the investigation of its behavior. Next we distinguish *optimization* (5), either of existing artifacts or of technical processes. This forms the last empirical project category. Other empirical goals are pursued in, for example, *proof-of-concept* projects, which we consider sub-goals of a design or design knowledge project, or in *exploration*, which may be part of a descriptive or means-end knowledge project. In the same way we consider *feasibility studies* to be exercises in business economics. Our sixth and final category consists of *formal* projects that consider problems from *mathematics* or information technology (6), which focus on formal proofs or algorithms and strictly speaking do not require empirical research. Through the years, these distinctions have resulted in a classification of engineering projects that consists of six categories, which, as will argue, covers all engineering projects in academic practice.

Figure 13.3 shows the distribution over the six categories of the sample of 82 projects carried out at Delft University of Technology between 2013 and 2015.

As can be seen, the first three categories cover the large majority of projects and the last three even cover together fewer projects than the smallest of the first three categories. Nevertheless, we include them as categories because they identify specific engineering activities with their own characteristic methodologies. In the next sections we provide theoretical descriptions of the classes and illustrate them with some examples from BEPs carried out in Delft.

13.3 Descriptive Knowledge Projects (1)

As mentioned before, the answer to the question whether an engineering project belongs to some category depends on the project's overall goal and consequently on the overall method used to achieve this goal. If this goal is descriptive knowledge, the project is a descriptive knowledge project. What we mean here by descriptive knowledge is the most traditional account of knowledge. It is the kind of knowledge expressed in sentences from natural or formal languages, which describe some specific natural, social or artificial feature of the world. Purely descriptive knowledge in engineering projects comes closest to the traditional notion of knowledge produced in the natural sciences. The standard methodology handbooks mainly discuss methods for achieving this goal. Consequently, in retrospect the process through which this knowledge is produced can be reconstructed along the lines of the well-known empirical cycle. Let us, for comparative reasons, summarize de Groot's (1969) version of the cycle, which consists of the phases of observation, induction, deduction, testing and evaluation.

In the observation phase, the empirical evidence is collected and grouped with the help of hypotheses that are already considered. Then, in the inductive phase, a specific hypothesis is formulated, preferably with (in)dependent variables that are well-defined so that the hypothesis can be tested against empirical data. Next, in the third, deductive phase, specific predictions are deduced from the hypothesis that can be empirically tested. In the fourth phase of the cycle, the empirical data are obtained or collected, often through laboratory experiments. Finally, in the fifth evaluation phase the empirical data, the outcomes of the experiments, are compared with the predictions from the hypothesis and it is evaluated to what extent the hypothesis did withstand the trial of empirical testing. It should be emphasized that this model is only a retrospective reconstruction of what happens in real research practice.

Not all projects that belong to this category complete the entire empirical cycle. When investigating completely new and unexpected phenomena, researchers often claim to do explorative research, which does not get as far as the testing of a hypothesis. We consider explorative research to complete only the first half of the empirical cycle. Consequently, it should result in well-formulated and empirically testable hypotheses.

Hypotheses that survive several empirical-cycle iterations become candidates for scientific knowledge. These products of descriptive research are traditionally characterized as structural descriptions of the world, which are as value-independent as possible (not only in the natural and engineering sciences but also in the social ones). The final products are theories or laws, which may be considered to have a definite truth value, and the appropriate method to reach these products is through inductive and deductive reasoning. From a societal point of view, these products should be freely accessible to the public. Of course we are well aware that all these claims are idealized and strongly contested in the literature, but we use them nevertheless to contrast them with the final products of design projects.

In engineering projects, purely descriptive knowledge differs from knowledge produced in the natural sciences in two important respects. First, in the natural and social sciences the goal is usually to produce knowledge that is as context-independent as possible; the more abstract, the better. Since engineering knowledge should serve the purpose of creating artifacts, it is inherently context-related and abstractness is less valued than in the natural sciences (Vries 2010). Second, descriptive engineering knowledge serves the purpose of creating artifacts, whereas knowledge in the natural sciences is collected for its own sake.

To illustrate the descriptive engineering knowledge category, let us consider the example of a descriptive knowledge project about hopping and rolling objects. The project leaders wanted to find out whether elliptical shapes rolling down a slope "reach higher steady state velocities than circular shapes". Under these circumstances elliptical objects start to "hop" and temporarily lose contact with the slope. As they are in the air, the rolling friction is zero. Consequently, the researchers hypothesized that an elliptic tube rolls faster down a slope than a circular one. They measured the various velocities and falsified their hypothesis, concluding that elliptical shapes do not roll faster than circular shapes. This is a perfect example of a descriptive knowledge project, and is similar to research projects in physics, although the latter are usually less directed at application.

13.4 Design Projects (2)

No category of engineering project differs more, with respect to the goal pursued, from descriptive knowledge projects than engineering *design* projects. The goal of projects of this category is the design of technical artifacts or processes that fulfill a set of specified design criteria, so that they behave or function according to pre-specified norms, something of a completely different nature than producing an approximately true description of aspects of the (artificial) world.

Design is an ambiguous term. It may refer to a process ('to' design) or a product ('a' design). As our classification hinges on the goals of the projects, we will focus on the latter. As a product term, design may directly refer to an artifact or process, or it may refer to a blueprint describing the construction of an artifact or process. A conceptual design is a generic solution to a design problem and forms a suitable starting point for the prototype design cycle, details included. Often, these conceptual designs are expressed in technical diagrams or sketches. A prototype is a physically functioning pre-phase of the artifact that leaves many details still open. Prototype designs enable designers to test the most important requirements of the design specification. During a prototype design cycle, the prototype is constructed (the synthesis) and tested (the evaluation) to find out to what extent it fulfills the design requirements. Conceptual and prototype designs both belong to the category of design projects.

To what extent does the aim of a design project determine the method that is to be used? Design methodology is less univocal than the methodology for obtaining descriptive knowledge. Some designers even deny the existence of general rules for solving design problems. This seems to be due to the context-relativeness and the uniqueness of a design. Nevertheless, in retrospect, many design processes can be reconstructed as going through several distinctive phases.[3] Many classifications exist, and ours is inspired by Jones (1992), Roozenburg and Eekels (1995), and Cross (2008).

To start with, in a diverging *analysis* phase, the functional design brief is analyzed, redefined, and perhaps divided into subproblems; the design brief is then fixed, the goals are set and the design specifications are operationalized. Secondly, in a transforming *synthetic* phase, the working principle is chosen and the design is constructed out of its parts; this results in a provisional design, or a prototype. Thirdly, in the *simulation* phase, the design team finds out through reasoning or experimentation to what extent the concept or prototype shows the required behavior. Finally, in the *evaluation* phase, it is decided, often by means of an evaluation matrix, whether the prototype satisfies the requirements sufficiently, and whether it requires further optimization or needs to be replaced by a new design proposal. Often the latter holds and then the design cycle starts again until the outcome satisfies the design requirements. To be sure, this representation is schematic and simplified. For instance, the synthesis phase may lead to new questions concerning the design requirements. Also, designers learn about the problem while working on the solutions.

Just as with the empirical cycle, some projects may complete only a partial design cycle. A proof of concept, for instance, may result in a well-functioning prototype that instantiates the concept or operational principle. If the goal is to assess the consequences of some new causal intervention, we consider a proof of concept to be the end of the first prototype design cycle iteration. The problem has been identified and a possible solution for causal intervention has been proposed and carried out. During the evaluation phase, it should become apparent whether the intervention has worked out as planned. Thus, a proof of concept may also be interpreted as a way to underpin some piece of engineering design knowledge or a 'principle of operation'.

The outcomes of a design cycle, such as a plan or a constructed artifact, are of a categorically different nature compared to the outcomes of an empirical cycle. In contrast to the latter, they are not true or false but (in)efficient or (in)appropriate. Moreover, being functional and intentional objects for societal use, they are inherently normative and value-laden. Next, the accompanying method is not inductive and deductive reasoning about hypotheses, but comes down to proposing possible constructions and evaluating them against the design constraints formulated at the start of the project. Design, in contrast to scientific knowledge, is generally strongly context-dependent. And the ways in which these products are manufactured are

---

[3] See ISO (2006) section 5, or the ABET (1988) definition of design, which states: 'Among the fundamental elements of the design process are the establishment of objectives and criteria, synthesis, analysis, construction, testing and evaluation.'

often protected by patents, whereas the outcomes of experimental science are, or at least should be, freely accessible to the public.

Let us turn to an example of the design category. It concerns a project in which the designers focused on the conceptual design of an underwater swarm vehicle (USV). It had to be no longer than half a meter and to be able to establish a swarm in North Sea water. It had to allow for an interchangeable payload and an individual USV had to be able to create, jointly with many others, a sensor network for communication, and it had to allow for the provision of sufficient energy for doing this. USVs are meant for underwater tasks such as the inspection of oil rigs, and therefore it had to be highly maneuverable. Decentralized swarm control was preferred, to avoid computational problems when the swarm increases. This decentralized swarm control is what made the project innovative. Finally, the USV to be designed had to be inexpensive (less than 1500 euros) to allow for production in large numbers. The execution of the project started with a functional decomposition of a USV, and questions about the working principle, communication, buoyancy, battery, and the charging procedure were analyzed. The designers chose a four-finned propulsion concept with six degrees of freedom. Their proposal required some experimenting regarding the dimensions of the fins and its actuators. The energy source was chosen to be electricity stored in on-board batteries. To charge them and to communicate the data, the USVs were equipped with a hydrophone to locate a base station, while mutual communication was guaranteed by phototransistors and pulsating blue light power LEDs. At the end of the project, no physical prototype or model was realized to find out the real (swarm) behavior of the USV's; the communication system still had to be tested, the electronics had to be integrated, and even the four-fin hydrodynamics had to be further investigated to come to a working USV. Still, since the final goal of this project was a working artifact, it is an obvious example of an innovative engineering design project.

13.5 Engineering Means-End Knowledge Projects (3)

Many engineering endeavors aim at producing and formulating means-end knowledge that prescribes how to act (and what to avoid) in order to achieve some prespecified goal in a specific context. All innovative engineering projects occupied with these endeavors belong to our third category of *engineering means-end knowledge*. This type of knowledge is often explicated as heuristics, rules of thumb, guidelines or protocols, which are then used in the design and maintenance of an artifact in some specific context. It suggests how to accomplish a predefined aim that does not yet obtain. Surely, not all know-how can be made explicit, but in engineering at least some of it can, although much engineering tacit knowledge seems to escape attempts to be spelled out completely.

Although an important part of the projects we encountered were preoccupied with the production of engineering means-end knowledge, strangely enough, we were unable to find any general methodological publications about this category of knowledge production and its accompanying methodology. Here, we refer to Niiniluoto (1993), where the author discusses the notion of a "technical norm", which originates in the writings of von Wright (1963a, b) on practical inference. Niiniluoto's (1993) formulation of a technical norm reads:

(*) If you want *A*, and you believe that you are in a situation *B*, then you ought to do *X*.

Leaving many questions and points of philosophical discussion regarding means-end knowledge aside, for the time being we go along with Niiniluoto, who writes: "we may propose that technical norms of the type (*) express the typical structure or logical form of the knowledge provided by design science" (Niiniluoto 1993, p. 12). Consequently, we take (*) to be the typical structure or the *canonical form* of engineering means-ends knowledge.[4]

Although means-end knowledge is about empirical interventions and should be argued for or underpinned by empirical means, a general means-end research methodology has not been developed yet. No handbook about engineering means-end knowledge exists today. According to Niiniluoto (1993, p. 13), engineering means-end knowledge may be validated or 'supported' top-down, by way of scientific theories or laws, or bottom-up using experimental methods if no appropriate scientific theories are available. Similarly, within the field of *Information Systems*, Hevner (2007) lets the design cycle be accompanied by a 'rigor cycle', which involves scientific theories, and a 'relevance cycle', which introduces the context into the design research. Theoretical ways to underpin engineering means-end knowledge differ from empirical ones. The first are based on knowledge about physical (im)possibilities and causal interventions. The second, in contrast, uses primarily trial-and-error methods, in which reliable artifacts are considered proofs of the reliability of some form of means-end knowledge. Interesting questions are whether trust in engineering means-end knowledge is based on induction, on knowledge about causal interactions, or on both, and through what methodology this means-end knowledge has been established. We have to leave these and related questions for future research.

Though a means-end design research methodology still has to be developed, we are convinced that the outcomes of this type of research differ in many respects from the results of the empirical cycle. These outcomes are not hypotheses based on (in)dependent variables. Designers' means-end knowledge of the form (*) is advice on how to produce an artifact with pre-specified characteristics within some specific context. This kind of design knowledge is empirical and should be grounded in reality to become more reliable. Prescriptions may turn out to be *unreliable or ineffective* (or "false") if an artifact fails to have the desired characteristics although the prescriptions were followed when the artifact was produced. Furthermore, in contrast to knowledge in the descriptive natural sciences, the resulting engineering

---

[4] The *X* in (*)'s consequence may be interpreted to be a sufficient, a necessary, or a rational means to achieve *A*. The consequence may also have different forms, such as doing *X* is effective, which would give (*) a more descriptive ring. We will leave these more philosophical subjects for another occasion.

means-end knowledge is inherently normative as it is related to action. It prescribes what is good to do in certain circumstances. Finally, the context-relatedness of means-end knowledge renders it highly non-monotonic. Since the description of the relevant context is open and never complete, introducing new information may render a prescription that is generally effective useless in some contexts. Normally, it is a good advice to extinguish a fire with water. If, however, we add that the burning object is made of a combustible metal, this advice becomes useless and may even be dangerous. Although some compare ceteris paribus conditions to the openness of means-end knowledge, descriptive knowledge of the natural sciences lacks this profound action-related non-monotonicity.

Just as proofs of concept are often partial design or means-end knowledge projects, normally, feasibility studies turn out to belong to an overall means-end knowledge project. These studies are always context-dependent and therefore profoundly non-monotonic. Additionally, if feasibility studies are generalized, they usually have a means-end form. If some device has to be produced (end) within the context of specific facilities and one has only a restricted budget at one's disposal (context), only some specific technology will be feasible (means).

We illustrate the engineering means-end knowledge category with a project that compares the energy-efficiency of gain scheduling with feedback linearization used to actuate and control magnetic manipulation systems. Since gain scheduling explores nonlinearities in magnetic fields better than linearization controlling, the researchers expected that gain scheduling was more efficient for magnetically manipulation of mechanical systems. To prove this idea they built an experimental comparative setup, and a simulation. The results of the experiment showed that with approximately the same settling times, the control effort of a gain scheduling controller was 7.5 times less than that of feedback linearization.

In their abstract the authors claimed: "It is found that a controller based on gain scheduling can perform the same reference trajectory with the same settling times using less control effort than a feedback linearization controller." At first sight, the 'that-clause' may be read as a hypothesis. Modal claims, however, make bad hypotheses. They are far too weak from a logical perspective. Claims that something *can be* the case do not form the basis of science. Moreover, positive modal claims cannot be falsified, just as negative modal claims escape verification. In this case, the modal claim suggests a practical orientation of the project. Modal claims articulate what engineers can, and even, should do in certain circumstances. The project is a contribution to engineering means-end knowledge and should be read as follows: "If you want (*G*) to actuate and control with the same settling time as a feedback linearization controller, but (*C*) should use less control effort, then (*M*) use a gain scheduling controller."

This means-end formulation of the result is stronger than the one actually stated, and perhaps the authors were not sure enough about their results to put them in the stronger means-end form. From a systematic point of view, however, it was knowledge in the means-end form that the project leaders were after.

13.6 Modeling Projects (4)

Often, engineers work with models. Even older engineering textbooks devote entire chapters to modeling and simulations (Krick 1969). If the production of an innovative model is the main object of an engineering project, we classify it as a modeling project. Engineers use models and modeling techniques as least as much as their colleagues in the natural and the social sciences. We take engineering models to be approximate representations of aspects of the real world, which in the case of engineering projects is the target system. Moreover, we assume that models are constructed and used for an explicit goal, which is not necessarily an epistemic one. Beside descriptive and means-end cognition, they may be used for exploration, comparison, optimization, simulation, construction, control, making decisions, and so forth. Engineering models serve many goals and come in many different forms. One way to order them is by the ways in which they are ‘materialized’. We consider five types.

First, models can be sets of interpreted (differential) equations. These sets are called mathematical models. They are approximate *mathematical* description of dynamic natural, artificial, or social phenomena. Examples are the Lotka-Volterra equations, Lorenz model of the atmosphere, Shockley diode equation etc. Their main concern is often to describe observational phenomena. In contrast to mathematical models, *theoretical* models focus on theoretical structures of the world, which are typically compared to a specific better known (observable) mechanism. Bohr’s model of the atom and the double-helix model of DNA are examples. Notice that both types of models are common in science and engineering. They may serve well in (molecular) descriptive and means-end knowledge projects.

Models less often used in the natural sciences are *iconic* ones, also called replicas and analogue machines, which are form reproductions for some purpose. They are “designed to reproduce as faithfully as possible in some new medium the structure or web of relationships in an original” (Black 1962, p. 222). This category comprises pictures, sketches, maps, dummies, mock-ups, prototypes – to study, for example, the aesthetics of a design – and scale models. Clearly, this kind of model occurs often in engineering practices which concern design and means-end knowledge. Engineers also use many forms of *structural* models, such as stream diagrams, electronic diagrams and function-block diagrams. Structural models do not copy some existing web of relationships in an original, but they describe and help to build the structures of a newly conceived technical artifact or process. Finally, *logical or axiomatic* models are the mathematical structures that verify an axiomatic system. They may occur in the mathematical project type considered below.

Since models appear in so many forms, what can be said about the method that accompanies the modeling category? The answer to this question becomes clear when one realizes that models are special kinds of artifacts or tools to achieve some goal. Thus, the accompanying method comes close to that of the design cycle, including a design brief, design requirements and an evaluation matrix for the final phase, but completed with standard criteria of model quality. A modeling cycle starts with a clear identification of the purpose or goal of the model, because this goal determines the (ir)relevant aspects of reality and the required accuracy. Taking into account the purpose and the context of use, modelers can then fix the design brief and the design requirements for the model. It should be clear when the model is successful and when it fails to achieve its purpose.

The different types of goals mentioned above imply many quality constraints for models, among which validity is an important one. Whether the aim is knowledge, simulation, construction, decisions making or control, the behavior of the model should at least partially correspond to that of the target system, although for some explorative purposes this may be less important. Other general constraints are, for instance, verification (absence of errors of calculation) and robustness (small effects of perturbations), and additionally for computational models complexity and efficiency.

Close relatives of the model types just mentioned and often encountered in engineering practice are *simulations* and *serious games*. Simulations may be viewed as huge number-crunching projects in which numerical solutions to differential equations are produced. They may also be interpreted as stand-ins for real-world systems on which we cannot carry out experiments, or even as a new type of science between experimental and theoretical (Winsberg 2010, p. 31). In any way, they at least comprise implemented mathematical models simulating dynamical behavior within a computational environment. Simulation projects are therefore considered modeling projects.

If engineers do not have sufficient information about how human actors will behave under certain technical circumstances, often they introduce real people into their models and let them occupy various roles in a serious game. These games may have various goals. An important one, besides teaching, is getting to know people's interactive reaction patterns while performing various roles; another is to predict how certain designs of socio-technical system will develop under specific circumstances. Regarding the former goal, serious games are instruments for social and behavioral scientists, whereas the latter goal is the more technological one. In the latter sense serious games can be considered as models of socio-technical target systems that comprise human beings occupying competing roles. Consequently, the development of simulations and serious games should and do exhibit the same type of design cycle as other models.

Let us turn to the modeling example, which concerns the *Robird Peregrine Falcon*. The engineering company *Clear Flight Solutions* had developed a flying robot made to look like a peregrine falcon to scare off birds that cause inconveniences. Skilled pilots are still needed to steer the flying robot from the ground. In this project the researchers tried to develop mathematical models to describe the gliding and the flapping flight of the robot. It was assumed that, besides forward thrust, gliding flight can be described by the same equations as flapping flight but with different parameter values. The models related three-dimensional aerodynamic forces and moments due to lift and drag to translational and angular accelerations in three directions. They covered the variables: roll; pitch and yaw (and their derivatives); velocities in the x, y and z directions; left and right aileron angles; tail deflection; flapping frequency. In addition to these variables, the models needed many parameters. Some of these (total mass and moments of inertia, the moment arms and areas, and lift/drag coefficients) could be (indirectly) determined. Others were approximated by parameter optimization. Examples were: the self-damping terms of the roll and of the yaw; with respect to the tail deflection the derivative of tail lift coefficient and the second derivative of tail drag coefficient; the moment arm rudder and the derivative of rudder force coefficient with respect to left and right aileron angles.

The parameter optimization was planned to be carried out by comparing simulated outcomes of the model based on a training dataset with a test dataset. Unfortunately, the researchers had to do without the Robird dataset and instead had to rely on the data of a model airplane. An additional problem was that the MATLAB 'pattern search' function was too slow to calculate all parameter values at once; moreover, the solutions did not seem feasible. To overcome this problem, the researchers first optimized the tail related parameters, and after fixing them, they turned to aileron parameters. They also estimated the interdependencies. The resulting model produced acceptable results regarding the yaw, but had difficulties with correctly predicting the roll and pitch of the model airplane.

Reading the report on the Robird peregrine falcon project gives the impression that the goal of the mathematical model is twofold: first it is descriptive knowledge of the relation between the different variables and parameters, but, secondly, the reason for wanting to know this descriptive knowledge is to support autopilot design, a goal related to control. Although the researchers did not go into setting up a design brief for the model, they did take seriously its validation, one of the most important design requirements for knowledge-geared modeling projects. They formulated it explicitly: the Variance Accounted For (VAR) of the model's predictions had to be larger than $80\,\%$ and the Normalized Root Mean Square Error less than 0.1. Although the goal of the model was not clearly stated and the model failed the validity criteria for pitch and roll, this does not invalidate the fact that it is a clear example of a modeling project.

13.7 Technical Optimization Projects (5)

The final empirical category of our classification concerns optimization. In the broadest sense of the term, engineering optimization projects aim at improving existing technical artifacts (also models) and processes. They are close in nature to redesign projects but since their method is distinctly different, optimizations establish an autonomous category. The outcomes of these projects are improvements.

Before turning to a general description of the method, we draw the reader's attention to the distinction between engineering and mathematical optimization. Many books and papers on optimization have been published. However, most of those are dedicated to *mathematical* optimization. Under the assumption that the problem is completely paraphrased by mathematical equations, the main subject is to find the mathematical optimum in the solution space. Although this is an important subject, here we would rather like to embark on the more general subject of engineering optimization.

In our terminology, engineering optimization can be achieved in two ways. First, engineers may improve a type of artifact or process by changing or adapting the operational principle of the artifact or of one of its main subsystems. For instance, the performance of a carburetor equipped internal combustion engine is improved by fuel injection. Secondly, engineering improvements may also be accomplished by changing the values of the design variables so that the results score better on the optimization criterion under the same constraints. They often achieve such improvement by applying mathematical optimization techniques.

Optimization projects evolve through the following phases, which together we call the *optimization progression*. First, the optimization problem should be identified. It should be clear which type of artifact or process with what aims is considered. Second, information should be gathered about the working principles applied and about the conditions and context under which the artifact operates normally. Moreover, the *optimization criterion*, i.e., the property of the artifact that should be improved, should be identified. Additionally, it should be clear how the traditional artifact scores on this criterion, and what change would count as an improvement. Third, the data collection will lead to the identification of the *design variables*, which are the variables that are freely adjustable during the design and the remaining variables or parameters that are constant or are to be kept constant. Additionally, the constraints governing the optimization should become clear. These are the *restrictions* that the optimization should fulfill and that cannot be altered. Fourth, the engineers should decide whether to improve by installing a new working principle or to stick to the old one and apply mathematical optimization. The latter involve modeling the system, which will take most of the time, and feeding the model into an optimization program to find an improved working point in the solution space. Finally, it should be evaluated whether the proposed change is really an optimization in comparison to the artifact's original performance.

Whether mathematical optimization is sufficiently innovative to be considered an innovative engineering project may be a point of discussion. Optimization plays an important role in many engineering problem solving activities. Take, for instance, the design cycle, where the successful first iterations may result in a proof of concept, after which the prototype is optimized and refined during the following iterations. Additionally, the model development in the previous example uses parameter optimization. Optimizations also play an important role in other context than the design cycle, for example in product development. The mathematical optimization algorithms only improve a well-known artifact configuration and will never propose to change its working principle. A global optimal design hinges on the choice of working principle and mathematical optimization (Parkinson et al. 2013, p. 15). Consequently, we consider projects in which an artifact is partly redesigned through introducing a new working principle, typical examples of the optimization category. Whether some mathematical improvement project is sufficiently innovative to be part of this category depends on the specific circumstances.

Our *example* of an optimization project concerns the ‘Selection, Modeling and Optimization of an Electric Motor for DUT Student Formula Racing’. The goal of this project was first to underpin the design choice for the off-the-shelf electric motor of the electric racing car developed by students at Delft University of Technology, and then to optimize this motor for this specific vehicle. The optimization criteria were power density, efficiency, braking power, heat production, and the torque-speed curve of the motor.

The DUT student racing team selected a three-phase, alternating current (AC), permanent-magnet synchronous electric motor (PMSM), for the propulsion of their vehicle. They took an AC motor since its torque is more constant than that of DC motors at lower speeds and decreases at higher speeds when the tire friction decreases as well. A three-phase motor was preferred because of its higher efficiency and favorable torque-current characteristics in comparison to a single-phase motor. Finally, a permanent magnet in the rotor was preferred to an inductive one because the most important current in latter runs through the rotor and in the former through the stator, which is easier to cool than the rotor. Moreover, a permanent magnet motor is more efficient than an induction magnet one, and it has higher dynamic response due to its higher power density and torque-inertia ratio (Tong 2014, p. 42).

Although the AMK© off-the-shelf servo motor brought the DUT racing-team many successes (DUT racing 2015), the designers wanted an optimized motor for the specifics of the complete powertrain, the body of the car, and competition regulations. To accomplish this task the project team made two models of the AMK motor as a reference. First, to study the sensitivity of the parameters the team members built an analytical model in MATLAB that used the Carter factor to correct for the air gap, but ignored all other losses and correcting factors. This reduced accuracy model reported a torque that was unrealistically high. Secondly, to predict the torque more accurately the team developed a finite element analysis model, using COMSOL’s magnetic-fields-physics interface. It had the disadvantage of being less clear about the inter-parameter relations. Although the model failed to predict the correct dynamic torque-speed curve, it did determine the correct static torque. Because modeling the motor was a substantial challenge, the team did not accomplish a real optimization proposal.

The electric-motor project illustrates the importance of modeling in engineering optimization projects. Yet, the project is not a pure modeling project because its main goal was to optimize an artifact (the electric motor of the racing car) and not to gain descriptive knowledge about PMS motors. The final goal was to find an optimum in the mathematical problem space, and to apply and test it in the real world. Moreover, alternative working principles were discussed and compared as well.

13.8 Mathematical Projects (6)

The sixth and final category of our classification covers projects that result in formal or mathematical objects. The two formal products that engineers most often seek to realize are either solutions to (applied) mathematical problems or algorithms and software. Regarding the first, the formal category covers all results or proofs in engineering mathematics. These results may concern claims in real and complex analysis, linear algebra, (systems of) differential equations, many forms of approximation theory (optimization!), Fourier analysis, Laplace transforms, applied probability theory, and so forth. Regarding algorithms and software, we consider most work in information science to be engineering knowledge projects, except the abstract branches of theoretical computer science, which would, however, still be projects in the formal knowledge category (Zill et al. 2011).

To attain purely mathematical results like formal proofs of theorems or approximation methods, engineers apply methods similar to those of mathematicians, which differ from the cycles we have seen applied in previous categories. George Pólya provides a worthwhile description of four different phases of mathematical problem solving. According to Pólya, first, the problem should be well understood and it has to be clear what is required exactly: what the assumptions are and what the goal is that must be reached. Next, by exploring the relation between the various items, the problem solver should design a plan about how to achieve her goal. After that, the plan should be carried out, which is often easier than creating it. Finally, in retrospect the solution should again be reviewed and deliberated (Pólya 2014, pp. 5–16). In the same publication Pólya formulates many useful heuristics for every individual phase.

We will not embark here on the philosophical discussion regarding the ontology of software and whether or not a program is an artifact.[5] We only observe that software development seems to follow the design cycle closer than Pólya's methodology, which is followed in the development of algorithms. Software development includes problem identification and requirements definition, which is followed by the design and construction of the software. Next, the software has to be tested and debugged and the process ends with the deployment and maintenance. Many standard software development methodologies exist, such as the Waterfall model, the V-model, the Spiral model, and so forth, which cover these or similar phases. Beside the validation and verification, software should comply with other standard constraints such as efficiency in time and memory, robustness, and fault-tolerance. Since the assessment of program robustness may be hard, it can become the goal of an entire project, which we will then classify as a subatomic mathematical project.

---

[5] Implemented software on a computer may be considered to be an artifact, and as such being part of a (software) design cycle.

Table 13.1 The six categories with their goals and accompanying method

	Category	Goal	Method
1	Descriptive knowledge	True world descriptions	Empirical cycle
2	Design	Successful artifacts	Design cycle
3	Means-end knowledge	Reliable, efficient heuristics	Theory and practice support?
4	Modeling	Models with specific goals	Modeling cycle
5	Engineering optimization	Specified optimizations	Optimization progression
6	Mathematics	Mathematical claims and objects	'How-to-solve-it' heuristics

The same holds, mutatis mutandis, for validation and verification projects, such as for instance the assessment and validation of a numerical solver.[6]

Our example of a mathematical project concerns a project with the title 'Obtaining design knowledge for an underwater robot-swarm for inspecting mooring chains'. Unfortunately, the project title fails to cover the project contents correctly. To prove this claim we quote the abstract in full length: "We propose a swarming-algorithm that enables a group of autonomous robots to inspect the mooring chains of an oil-platform located in the North Sea. The algorithm, from which an autonomous underwater robot swarm can emerge, is designed to work within the limitations of underwater robotics and communication. Agent-based modelling and simulation (ABMS) has been used to assess the influence of communication parameters and swarm-size on swarm behavior in order to obtain knowledge about underwater robot-swarm design. Using the gained knowledge an algorithm is constructed for an autonomous underwater swarm where 160 agents can visit and inspect at least 80 % of the mooring chains of a specific oil-platform, within 24 hours, without losing agents." This abstract clearly shows that the project's deliverable is not design knowledge, but an algorithm, although its assessment by the standard algorithm criteria was left to another occasion.

Table 13.1 depicts our categories with their defining goals and methods.

13.9 Profiles of Multiple-Purpose Projects

In the previous sections, we have introduced our six-type engineering project classification, which covers subatomic, atomic and molecular engineering projects. Projects with a single overall goal and accompanying method are atomic. If a project's goal or method is only part of a goal or method defining a category, the project is subatomic. If a molecular project combines different subgoals that fit our categories into an overall main goal, this main goal determines the project's category. Note that especially molecular projects may be composed in many different ways.

---

[6] In fact we did encounter these projects within the Delft Applied Physics department, which illustrates the differences of emphasis between the departments in a university of technology.

Molecular projects may have very different means-end hierarchies underlying the method that accompanies its main goal. The combinatorial possibilities are, at least in theory, almost limitless. Any type of goal can be served by any combination of subgoals and methods from the same or other categories. Typically, a design project requires several optimization iterations; descriptive knowledge acquisition normally involves the design of experiments and equipment; optimization calls for modeling and characterization of existing artifacts; and means-end knowledge scarcely goes without descriptive knowledge and additional know-how; and so forth. Before the introduction of the systematization of engineering projects presented in this chapter we did not have the means to study empirically the possible regularities within the vast realm of molecular engineering projects.

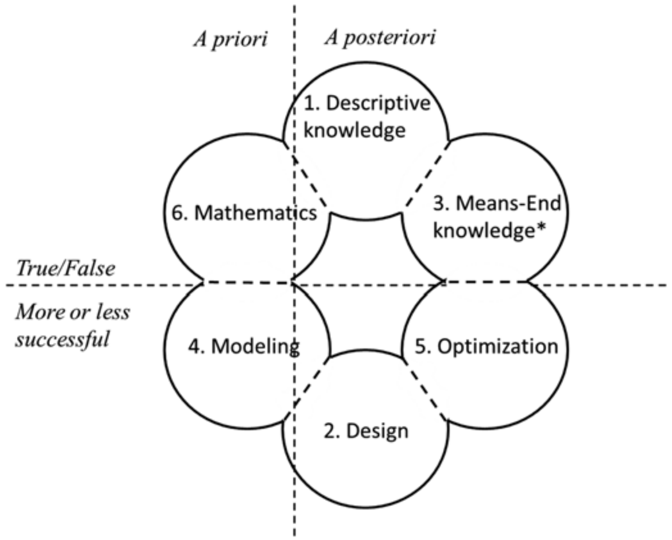
This brings us to the innovative engineering projects that have *more than one* overall purpose, which we refer to as multiple-purpose projects. Although these purposes fail to serve one another completely, they often partially share some subprojects. An experiment, for instance, may serve more than one purpose, such as a technical and a theoretical one. While an experiment may establish descriptive knowledge about the correlation between independent and dependent variables, it may also serve the goal of finding out whether, and if so, how a causal intervention may result in some desirable effect. For instance, experiments that establish knowledge about granulation of bacteria may be used for technological and theoretical biological purposes.[7]

As announced in the Sect. 13.2, we propose to characterize multiple-purposed projects with *Radar Charts*, which display the time spent on the main goals (and not the subordinated ones). The six axes of the charts represent our six categories. They are, in clockwise order: (1) Descriptive knowledge (North); (3) Means-end knowledge; (5) Optimization; (2) Design (South); (4) Modeling; (6) Mathematics. Note that due to this ordering, the outcomes of 'northern' category projects, 1, 3, 6, qualify for having truth values. Whether means-end knowledge should be considered true or false is still the subject of debate. Von Wright (1963b) refrains from attributing truth values to technical norms (Chap. 1, Sect. 1.7). Niiniluoto (1993) disagrees because without truth values, they are unable to convey knowledge. We think that means-end knowledge should be characterized as more, or less, reliable or helpful in some given context, and does not always have a definite truth-value. In contrast, the outcomes of projects from the 'southern' category are not true or false. They fulfill their requirements more, or less, successfully. Additionally, the north-south axis divides a-priori projects (5, 6) from the a posteriori ones. Figure 13.4 shows our six-category topology. The dotted lines dividing the categories indicate that for some projects their appropriate location in one of two neighboring categories is open for discussion.

By representing multiple-purposed projects in the way described, the resulting radar charts provide a *profile* or an "X-ray" of these projects through the time spent on its ultimate goals. In the scheme, half of the categories are related to knowledge;

---

[7] This example stems from the Nereda® wastewater treatment project, which is a clear illustration of a multiple-purposed project. For a description of this project see Zwart and Kroes (2015).



* Means-end knowledge is rather *more or less reliable* depending the context

Fig. 13.4 The topology of our categories

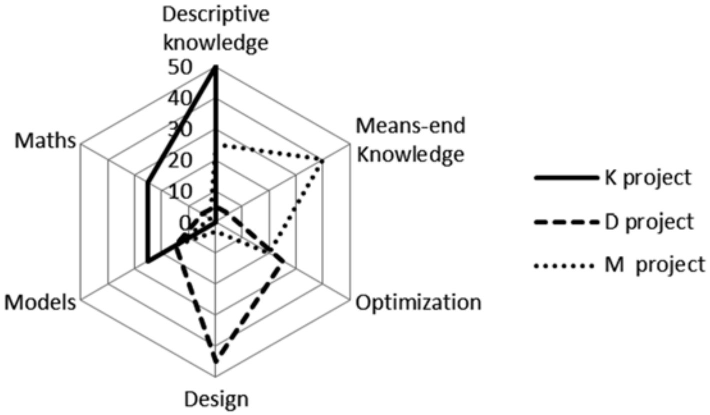


Fig. 13.5 Radar charts multiple-purposed Engineering projects

the other half concerns action. Within a radar chart this makes, for instance, "North-West" projects, i.e., projects where most of the time is spent in the North-West region, typical knowledge projects (K-project in Fig. 13.5), whereas "South-West" projects (D-project in Fig. 13.5) are mainly concerned with design and optimization. Represented in radar charts, our classification provides a useful instrument for multiple-purposed project representation and analysis.

Looking at the radar charts, one may wonder whether they do not also provide an appropriate way to characterize molecular projects. After all, a molecular project consists of a means-end hierarchy of atomic subprojects, which are all categorized along the six axes of our radar charts. Once this hierarchy is categorized, we can sum up the time spent along the axes. Although technically possible, we are doubtful whether this procedure would result in comparable summaries of molecular projects, because the information about how much time is spent at what location in the hierarchy is lost in the addition. Within an overall design project, many small mathematical subgoals deep down in the hierarchy may sum up to considerable time spent on mathematics. A radar chart of such a design project would give the impression that it is close to a mathematical project, whereas in fact it is not. Radar charts work primarily for multiple-purpose projects, because the main aims of such projects are unrelated.

13.10 Discussion and Outlook

Until now, we presented, on empirical grounds, a classification system for innovative engineering projects. The categories are defined by the main goals of the projects, considered as ends-in-themselves, and their complete accompanying methods, which together form a coherent unity (see Table 13.1). In this final section, we discuss the status of our classification and its philosophical relevance and we suggest some follow-up research.

Regarding the status, we start with the question whether the classification does not harbor *too many categories*, and ask whether categories are *independent*. If, for instance, the main goal of a (mathematical) model is an empirically adequate prediction of the (natural or artificial) world, it comes close to a descriptive-knowledge project. In such cases the underlying design cycle bears resemblance to the empirical cycle – even if the formulae used for prediction do not mirror the structure of the real world. Many other cases of similarity exist. For the construction of a scale model a method is applied that is much more similar to the design cycle than to the empirical cycle. Optimization projects may result in an improved design of some artifact type and therefore bear similarities to (re)design projects. Engineering design knowledge projects may also conclude with suggestions about how to improve an existing design. Although from a theoretical perspective these suggestions sound reasonable, in practice most engineering projects clearly fit one of our categories, because of the coherence between goal and method. Engineers go about it differently when building a mathematical model and when formulating and testing a hypothesis, although the model's validity is an important model-building requirement. Moreover, if modeling were a part of design, we would have to distinguish between design-projects-of-models and design-projects-of-non-models because of the differences in method.

To push the question further, we may also ask why optimization projects are not parts of design projects, since many of them cover prototype optimization. The reasons are, first, that in optimization projects the project team is not involved in the original artifact design, and second, that optimization projects have a definite own method. Moreover, although in practice many design projects integrate optimization activities, this is not necessarily the case, and optimization in design may be considered a separate subproject. The same does not hold for proof-of-concept projects, because the aims of these projects are not ends-in-themselves. They always serve a larger objective, as is the case, for instance, when proving how some operational principle can be successfully put to work in a design. Additionally, a proof of concept hardly has any distinctive accompanying method other than the application of the operational principle at issue. Consequently, it is more adequately categorized as a part of the first iteration of the design cycle, if the final goal is a design. Considering all specific aims and accompanying methods, we conclude that we need at least six different categories.

Granting that our six categories are independent still leaves open the possibility that we have *too few* categories and that we may have overlooked other important ones. At the Department of Applied Physics for instance, we encountered two other types of projects that might belong to categories not yet identified, namely *assessment of existing models* and *choice of materials with desired properties*. Why, however, do these not delineate independent categories? As for the proof-of-concept projects, we consider the assessment-of-existing-models projects to be only a part of atomic modeling projects, provided the project team is close enough to the modeling one. If not, the reason for the assessment will make such projects part of an engineering means-end-knowledge, or an optimization project. The former is the case if the overall goal of the project is to find out in what context the model is appropriately applied to achieve what goal, and the latter is the case if the objective is to improve the model. Projects that aim at a choice of materials with desired properties do not make a separate category, because this aim is never an end-in-itself. The properties are needed for achieving some overarching objective. Such projects might be part of an atomic means-end knowledge project if this ultimate goal is to find out what materials to use for which purposes under what circumstances. It may also be part of an atomic design project if its ultimate goal is a design. Moreover, a category for choosing materials with desired properties would lack a well-defined method, which makes it unsuitable for a separate category. To summarize, the additional consideration of projects conducted at the Department of Applied Physics did not result in new categories, which strengthens our belief that our classification is intra-academically complete.

This leads naturally to the question to what extent our classification is generalizable. Our search in the Department of Applied Physics for new types of projects revealed a difference of emphasis between this department and the Department of Mechanical Engineering. In Applied Physics, more projects are formulated in which some existing model, theory or method is validated, whereas in Mechanical Engineering more innovative design projects are carried out. This shows, by the way, that our classification is also a useful instrument for characterizing departmental differences within universities. Despite the differences in emphasis, we claim that our classification covers all projects within the two departments investigated: regarding these departments, it is intra-departmentally complete. The above generalizability question then reads: to what extent does our classification apply to the projects in other departments of technical universities? Let us call this empirical issue the intra-academic generalizability. We are quite optimistic about the answer to this question. Not only because the application of our classification to projects in the Department of Applied Physics went smoothly (a demanding test, given the differences between the disciplines), but also because of more theoretical reasons: to our knowledge we have considered all high-level-of-abstraction methodologies available. Consequently, we hypothesize that our classification is intra-academically complete.

Future research should reveal whether our hypothesis holds true. Two questions arise. First, it should be tested whether our categorization covers the projects of the same type of departments in other technical universities, both nationally and internationally. Second, it should be examined if the projects of departments we have not yet studied are covered as well. If our taxonomy is strong enough, it might be interesting to establish the bar-charts (cf. Sect. 13.2) of other departments. We have already seen that the charts of the Departments of Mechanics and Applied Physics in Delft are different. It would be interesting to test our classification in Departments of Industrial Design or Departments of Architecture, which are usually considered to be of a quite different nature. Another relevant intra-academic generalization question is whether we can discern specific types of molecular projects. That is, whether we can find recurrent hierarchical means-end patterns within the same types of goal-subgoal distributions. Still another is whether we can discern characteristic structures for the specific branches of engineering, or 'object worlds' in the words of Bucciarelli (1996).

The second major generalizability question concerns the *extra-academic* comparison of our classification system. By this we mean the question to what extent our classification applies to engineering in practice outside the university. Here the theoretical reasons are of less weight than for intra-academic generalizability. After all, it might be that the engineering method literature is primarily an intra-academic affair, which disregards engineering goals and methods used in practice. To assess the adequacy of engineering project education, follow-up research may focus on the extra-academic adequacy of our typology. We may want to find out whether structural differences exist between the ways academic and commercial partners go about while carrying out engineering projects. For instance, the engineering enterprises may prefer a more bottom-up approach (trial and error) whereas in academia engineers use more theory-driven (top-down) methods. Furthermore, it might be interesting to look for typical distinctive X-ray profiles in the different professional engineering branches.

Another generalizability-connected question asks whether our classification establishes a (mathematical) *partition*. If so, the categories have to be mutually exclusive and collectively complete. The collective completeness of our system has been discussed in the previous paragraphs. For mutual exclusiveness it suffices to consider the characterizations of the categories by having a quick glance at Table 13.1. The goals and their accompanying methods that define the six categories are so dissimilar that the demarcation between the categories does not leave open the possibility of an overlap, with perhaps 'models' being the only possible exception. After all, they are artifacts. Yet, we maintain that models, as approximate representations with a specific goal, are readily distinguishable from other artifacts. So, to make the nomenclature mathematically precise we should substitute 'artifacts' with 'artifacts that are not models'. We did not do so for practical reasons. However, with this substitution in place the six categories are mutually exclusive.[8] Another question is whether some projects may fall in more than one category. Empirically it turned out that most projects straightforwardly belong to one class of our categorization, because of the determining role of the goals and accompanying methods. Yet, we readily admit that for some projects the choice of a category may invoke discussion. This does not undermine the mutual exclusiveness of our categories but underlines the vagueness of some project descriptions, as they are not explicit enough about their overall goal and applied method.

Summarizing, we *claim intra-departmental completeness* of our classification regarding the Departments of Mechanics and Applied Physics. We underpin this claim using empirical and theoretical arguments. First, during 12 years among the almost thousand projects we encountered, very few could not be fitted into our categorization, and second, nearly all types of abstractly articulated engineering methods are covered by our classification. This convinces us that our categorization is not only valid for the two departments investigated but might hold within the other departments as well, so *we hypothesize intra-academic completeness* of our classification. It remains to be investigated, whether this hypothesis holds true for extra-academic engineering practices. We have good hope it does. To the best of our knowledge all general engineering aims and methods considered in the philosophy of engineering literature are covered by our classification system. So, we are eager to discover general engineering problem solving goals not covered by our system and encourage scholars and engineers to investigate the extra-academic completeness of our classification.

Now that we have discussed the characteristics of our categorizations, let us turn to some of its philosophical implications. As mentioned in the introduction, the empirical part of our work presented here responds to the call for an empirical turn by Kroes and Meijers (2000). Consequently, our research belongs to the empirical or practice turn in the philosophy of science and technology where scholars study the daily practices of scientists and engineers in order to reach empirically adequate descriptions of what they 'really' do (Soler et al. 2014). Overall, the difficulty with these studies is the descriptive-normative relation. When does the descriptive mode start to have normative impact, and how is this normative impact to be justified? Our classification is no exception to this rule. It is built on data gathered over a period of 12 years, and at the end of this period we started to assess project descriptions by means of the categorization we had arrived at, witness the example of the

---

[8] To push the artifact-argument further we could maintain that all the goals of our classification are artifacts in some sense. Accepting this counterargument would therefore block any categorization of the main engineering problem solving activities. Consequently, the counterargument may be easily parried within the context of the willingness to set up such classification.

mathematics category described in Sect. 13.8. Categorizing human activities always requires some readjustment or reconstruction.

We choose not to say much overall about the difficult question of the descriptive-normative relation in science and technology studies. In our specific case, however, the classification started with descriptions of the individual projects, covering their final goals and methods. After years of experience with students and supervisors struggling to describe adequately an engineering project description, we have arrived, by way of a series of less-satisfactory classifications, at our final categorization. In the end, the sheer number of projects that fit the final system started to suggest that students were well advised to formulate their projects according to our categorization. This is not the same, however, as drawing normative conclusions from descriptions. It only shows that human actions may successfully anticipate empirical regularities.

Let us return to the question of the differences between science and technology referred to in the introduction. With the emergence of telescopes, thermometers, barometers, and so forth, science became undeniably closely related to technology. Vice versa, with the use of geometry and mechanics, and later, with the introduction of thermodynamics and electrodynamics, technology became closely related to the sciences. Today science and technology are inextricably intertwined. Nevertheless the two are not the same and the differences between the two are more difficult to articulate than, for instance, the difference between 'owl' and 'beaver', as the terms 'science' and 'technology' are both abstract and lack direct referents. This causes their similarities and differences to depend on one's frame of reference or point of view. From a sociological perspective, for instance, one may study the practices and dynamics of individuals and groups of scientists and engineers. These practices and dynamics are quite similar, and so are science and technology from this point of view. People working in an R&D laboratory often have a hard time to tell whether they see themselves primarily scientists or engineers (Vries 2005). In the wake of Latour and Woolgar's *Laboratory Life* (1979), sociologists of science have carried out many empirical studies within laboratories. Latour has studied the daily activities of employees in the molecular biology laboratory of the San Diego Salk Institute. He and many sociologists of science inspired by him do not distinguish between science and technology and study the practices of 'techno-science.' From their perspective the difference between modern science and technology is a myth.

Another perspective is to study the goals and methods within science and technology as they appear in project proposals and reports. The sociological perspective may depict a more realistic picture of the human side of science and technology. Our means-end analysis, however, is more distinctive regarding the products of both enterprises, despite the rationalizations these reports undoubtedly contain, due to group dynamics and personal motives. Neither of these perspectives, however, gives a truer or a better picture of science and technology as they 'really are'. All we can say is that the sociological perspective seems more appropriate when the process is concerned and ours more for the products of science and technology. Inspired by Kroes and Meijers' call for an empirical turn, we have adopted the second perspective and have investigated the relation between science and technology through the aims and accompanying methods of the projects executed in these disciplines. We have chosen to study the differences and similarities between the two by painstakingly investigating the means-end relations within science and engineering project descriptions. Thus, we depict the relation between science and technology by using the contrasts between the goals and methods of innovative projects as they are sanctioned or rejected by the appropriate research agencies that shape the future of both enterprises. We may expect that in traditional physics projects, descriptive knowledge, modeling and mathematics are more emphasized than design, means-end knowledge and optimizing, but without empirical evidence this expectation remains speculation.

What has our research methodology brought us so far? We began by claiming that our research has proven the insufficiency of the technology-as-knowledge perspective. We encountered many projects in which the main goal was not to produce descriptive knowledge but rather the design of procedures and artifacts, trustworthy engineering means-end knowledge, well-functioning models, optimizations, and formal results in mathematics and computer science. Regarding our overarching goal, to shed light on the relation between science and technology, we need to observe that we still have to do without corresponding investigation within the sciences. For the reasons sketched above, the social studies of science did not yield a corresponding categorization of scientific projects within scientific practices. Without such a result, unfortunately we cannot assess the fruitfulness of our approach for the characterization of the science-technology relation. We are looking forward to a similar categorization of scientific projects in the future. Investigating BSc or MSc end projects in physics departments of traditional universities would make an interesting start for achieving such a categorization. This illustrates the fertility of our research methodology. Our taxonomy helps to conceive how to study and describe the various relational structures between engineering practices and different kinds of engineering knowledge. Similarly, it may even shed light upon the relation between theory and practice within the practical and the descriptive sciences and may have interesting implications for practical philosophy overall.

This brings us finally to the fulfilment of Kroes and Meijers' 'prophecy' that opening the black box would lead to new research topics. Due to the empirical turn we have discovered a large blind spot in the standard literature on methodology and philosophy of technology. Within these realms one of the most important products of universities of technology, namely engineering means-end knowledge, has remained almost entirely unexplored territory.[9] No integral answers are to be found in methodology handbooks to questions like what the logical form engineering means-end knowledge is, how this knowledge comes about, how it is to be tested, confirmed or falsified, and what its relation is to the descriptive knowledge of the other sciences. If, as a first result, these questions are now going to be addressed by the methodology community, methodologists, philosophers or engineers, the authors of this chapter will be more than satisfied with the fruits of their labor.

---

[9] Some isolated initiatives, however, can be found, for instance, in de Vries et al. (2013).

Acknowledgements The authors want to thank the editors of this volume for their efforts to improve the readability of the text.

References

Black, M. (1962). *Models and metaphors: Studies in language and philosophy*. Ithaca: Cornell University Press.

Bucciarelli, L. L. (1996). *Designing engineers* (1st ed.). Cambridge: The MIT Press.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks: Sage.

Cross, N. (2008). *Engineering design methods: Strategies for product design*. Chichester/Hoboken: Wiley.

de Graaff, E., & Kolmos, A. (2003). Characteristics of problem-based learning. *International Journal of Engineering Education, 19*(5), 657–662.

de Groot, A. D. (1969). *Methodology: Foundations of inference and research in the behavioral sciences*. The Hague: Mouton.

de Vries, M. J. (2005). *80 years of research at the Philips Natuurkundig Laboratorium 1914–1994*. Amsterdam: Pallas Publications (Amsterdam University Press).

de Vries, M. J. (2010). Engineering science as a "Discipline of the Particular"? Types of generalization in engineering sciences. In I. Poel & D. E. van de Goldberg (Eds.), *Philosophy and engineering: An emerging agenda* (pp. 83–94). Dordrecht: Springer.

de Vries, M. J., Hansson, S. O., & Meijers, A. (Eds.). (2013). *Norms in technology*. Dordrecht: Springer.

DUT Racing. (2015). In *Wikipedia, the free encyclopedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=DUT_Racing&oldid=674734805

Hevner, A. (2007). The three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), 87–92.

Houkes, W. (2009). The nature of technological knowledge. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 309–350). Amsterdam: Elsevier.

Hubka, V., & Eder, W. E. (1996). *Design science: Introduction to needs, scope and organization of Engineering design knowledge*. Berlin: Springer.

Jones, J. C. (1992). *Design methods*. New York: Van Nostrand Reinhold.

Koen, B. V. (2003). *Discussion of the method: Conducting the Engineer's approach to problem solving*. New York: Oxford University Press.

Krick, E. V. (1969). *An introduction to engineering and engineering design* (2nd ed.). New York: Wiley.

Kroes, P., & Meijers, A. (2000). *The empirical turn in the philosophy of technology*. Amsterdam/New York: JAI.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. Beverly Hills: Sage.

Layton, E. T. L., Jr. (1974). Technology as knowledge. *Technology and Culture, 15*, 31–41.

Meijers, A. (Ed.). (2009). *Philosophy of technology and engineering sciences*. Amsterdam/London/Boston: North Holland.

Niiniluoto, I. (1993). The aim and structure of applied research. *Erkenntnis, 38*(1), 1–21.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K.-H. (2007). *Engineering design: A systematic approach*. Berlin: Springer.

Parkinson, A. R., Balling, R., & Hedengren, J. D. (2013). *Optimization methods for Engineering design*. Provo: Brigham Young University.

Pólya, G. (2014). *How to solve it: A new aspect of mathematical method*. Princeton: Princeton University Press.

Roozenburg, N. F. M., & Eekels, J. (1995). *Product design: fundamentals and methods.* Chichester/New York: Wiley.

Sheppard, S., Colby, A., Macatangay, K., & Sullivan, W. (2007). What is engineering practice? *International Journal of Engineering Education, 22*(3), 429.

Soler, L., Zwart, S., Lynch, M., & Israel-Jost, V. (2014). *Science after the practice turn in the philosophy, history, and social studies of science*. London: Routledge.

Tong, W. (2014). *Mechanical design of electric motors*. Boca Raton: CRC Press.

Vincenti, W. (1990). *What Engineers know and how they know it: Analytical studies from aeronautical history*. Baltimore: Johns Hopkins University Press.

Von Wright, G. H. (1963a). Practical inference. *The Philosophical Review, 72*(2), 159–179.

Von Wright, G. H. (1963b). *Norm and action: A logical enquiry* (1st ed.). London: Routledge & Kegan Paul PLC.

Wilson, E. B. (1952). *An introduction to scientific research*. New York: McGraw-Hill.

Winsberg, E. (2010). *Science in the age of computer simulation*. Chicago: University of Chicago Press.

Zill, D., Wright, W. S., & Cullen, M. R. (2011). *Advanced Engineering mathematics*. London: Jones & Bartlett Learning.

Zwart, S. D., & Kroes, P. (2015). Substantive and procedural contexts of engineering design. In S. H. Christensen, C. Didier, A. Jamison, M. Meganck, C. Mitcham, & B. Newberry (Eds.), *Engineering identities, epistemologies and values* (pp. 381–400). Dordrecht: Springer.

Sjoerd D. Zwart is assistant professor at the Delft and Eindhoven Universities of Technology, The Netherlands. His PhD-thesis is about distance measures on Lindenbaum algebras (*Refined Verisimilitude*. Dordrecht, Kluwer, 2001). His research interests are: methodology and ethics of engineering and design. Most recently he has taken up studies into practices and tacit knowledge. Latest books: (2014) with Léna Soler, Michael Lynch and Vincent Israel-Jost: *Science after the Practice Turn in Philosophy, History, and Social Studies of Science*, Routledge; and (2015) with Ibo van de Poel and Lambèr Royakkers: *Moral Responsibility and the Problem of Many Hands*, Routledge.

Marc J. de Vries is an affiliate professor of Christian Philosophy (of technology) at Delft University of Technology, the Netherlands. He also has a chair in Science Education at the same institute. At Eindhoven University of Technology, the Netherlands, he serves as an assistant professor in philosophy of science and technology. His background is in physics (MSc). He is the editor-in-chief of the International Journal of Technology and Design Education, published an introductory book on the philosophy of technology for educators ("Teaching About Technology"), wrote a monograph on the history of the Philips Research organization and is a member of the editorial board of the Philosophia Reformata international journal. He co-authored the book The Matrix Code, on science fiction movies and Christian philosophy.

Chapter 14
For the Benefit of Humanity: Values in Micro, Meso, Macro, and Meta Levels in Engineering

Byron Newberry

Abstract The goal of this essay is to sketch a taxonomic outline of values within engineering. Any desire to understand the technology-society relationship would presumably benefit from investigating the values that inform engineering work, since that work is largely proximate to the production of technologies. More specifically, an understanding of the values constitutive of and operating within engineering at a multitude of levels can potentially aid in understanding how engineers go from thought to thing in the processes of design and manufacture. I propose a four-level hierarchy of engineering values, at the micro, meso, macro, and meta levels. Values at the micro level correspond to those values operative at the level of specific, detailed engineering tasks. Meso level values are those values operative in the process of translating functional descriptions of designs into structural descriptions – that is, at the creative level of engineering design. At the macro-level, I refer to the values operative for engineers at the economic/organizational level – that is, at the level at which engineers intersect heavily with non-technical interests. Finally, the meta level comprises overarching values that presumably inform all of engineering work.

Keywords Engineering values • Engineering design • Engineering organizations • Engineering profession

14.1 Introduction

Physicist Alvin Weinberg (1970), notable for coining the term *technological fix*, published a 1970 essay titled, “The Axiology of Science”, in which he explored values in science. While the article contributed to the general philosophical aim of better understanding the values undergirding judgments, appraisals, choices, and

---

B. Newberry (✉)
Baylor University, Waco, TX, USA
e-mail: Byron_Newberry@baylor.edu

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_14

priorities in science, Weinberg's objectives were ultimately more pragmatic. He sought to inform political and administrative discourse with respect to how scarce resources were allocated for scientific endeavors. Along the way, he sketched the beginnings of a taxonomy of scientific values, albeit as a science practitioner venturing into philosophical territory. As an example of an *implicit* value pervading science, Weinberg proposed the adage "pure is better than applied." That is, he perceived that many scientists took it as almost axiomatic that pursuing scientific knowledge for the sake of knowing was more intellectually desirable or satisfying than working on applications of existing scientific knowledge. As an example of an *explicit* value assessment, Weinberg noted that decisions about allocating resources to scientific endeavors were inevitably influenced by judgments about the "caliber of practitioners" in one place or another. In this case, judgments about the capabilities of certain research groups were made consciously based on weighing relevant information and experiences.

The goal of this essay, similar to Weinberg, is to attempt to sketch a taxonomic outline of engineering values, at least as seen from the perspective of an engineering educator-practitioner. Much has been written on the related topic of technological values, particularly as regards the interplay between the technological and social worlds. Less attention has been devoted to engineering values specifically. But, generally speaking, any desire to understand the technology-society relationship would presumably benefit from investigating the values that inform engineering work, since that work is largely proximate to the production of technologies. More specifically, an understanding of the values constitutive of and operating within engineering at a multitude of levels can potentially aid in understanding how engineers go from thought to thing in the processes of design and manufacture.

Sven Ove Hansson (2013), in discussing value statements applied to technology, differentiates between value statements applied at the macro level and those aimed at the micro level. Macro-level statements, according to Hansson, deal with assessments of technologies as wholes, or technology as a whole. Micro-level statements, on the other hand, are applied to individual parts or practices, such as whether one potential element of a design is better than an alternative. Hansson asks how valuations at the micro and macro levels are related, and suggests that we can gain insight from understanding such relationships. While Hansson focuses on values ascribed to technology, this essay will focus on values as expressed by engineers. The two are clearly related since engineers are the ones likely to express many of the values ascribed to technology, particularly at Hansson's micro level, but the two are not coterminous.

I will propose a four-level hierarchy of engineering values, at the micro, meso, macro, and meta levels. To help explain the boundaries of these levels as I will define them, I will first invoke the notion of the dual nature of technical artifacts (Kroes 2010, e.g.), which was elaborated to help distinguish between, and understand the relationship between, the structural and functional descriptions of artifacts. Valuations at the micro level, in my scheme, will correspond more or less to the structural side of things, to assessments of the bits and pieces that engineers work with. My meso level will then correspond to the functional side of things, to engineers' valuations relative to the specifications they are trying to satisfy. At the macro-level, I refer to the values operative for engineers at, for lack of a better term, the economic/organizational level; that is, at the level at which engineers intersect heavily with non-technical interests. Finally, at the meta level, I mean a level of overarching values. The title of this essay, for example, reflects a meta-level value espoused by most engineering professional organizations, a value that ostensibly underpins all of engineering work.

Taxonomies are constructions of non-essential categories for the sake of convenience, and such is the case for the hierarchy proposed above. The values expressed by engineers, or expressed within or about engineering endeavors, may be parsed in other ways, such as ethical, aesthetic, cultural, organizational, economic, technical, professional, and personal. As in Weinberg, some values may be thought of as implicit, meaning that they reflect prejudices so deeply held that they are hardly recognized as axiological, and rather tend to be taken as axiomatic. Other values are explicit, meaning that judgments are arrived at by more well-established criteria, or at least with more explicit reasoning. Hansson et al. (2013) suggest a division according to moral, legal, evaluative, and instrumental norms. Ibo van de Poel (2015) classifies values in engineering and technology in two ways, distinguishing between internal and external values, and between final and instrumental values. The former pair relates to whether a value important in engineering work arises primarily within the activity of engineering itself, or rather arises from the outside the activity. The latter pair relates to whether a value is intermediate to the achievement of some other value, or if it is pursued for its own sake. Van de Poel (2013) also proposed a hierarchy of values in an effort to understand how values get converted into design requirements in engineering. He posits a sequence in which general values get translated into prescriptive norms, which in turn get translated into specific requirements. Different methods of categorizing things – in this case, values – typically serve to highlight different aspects of those things, and differing classification schemes often necessarily cut across the boundaries of one another.

There are both descriptive and normative motivations for studying engineering values. If, as is widely argued, technological artifacts are not value-neutral, then understanding the values artifacts transmit and promote is of great interest, as is understanding how such values become embedded in those artifacts, a process in which engineering likely plays a significant role. So descriptively understanding the types and roles of values in engineering is important to gaining a complete picture of the technology-value dynamic. Further, engineering organizations are constantly crafting, whether explicitly or implicitly, the identity and ethos of the engineering profession, and engineering educational institutions are constantly inculcating an engineering mindset into the next generation of engineers. These are normative activities. Making clear and explicit the values that are being promoted can make for more informed discussion and debate about what values should be promoted.

Like Hansson and van de Poel, I wish to ask the question of how the values at different levels are related, and what we can learn from those relationships. Particularly of interest for the philosophy of engineering and technology, I think, are (i) the ways in which meta- and macro-level values impact meso-level values (that is, how the overarching values of engineering, along with the values expressed by engineering organizations and their stakeholders, impacts the values that inform the mapping of functional descriptions into the physical instantiations of artifacts); and (ii) the ways in which meso-level values and micro-level values interact (which may be germane to the non-unique mapping between structure and function).

In discussing my proposed taxonomy of values in the subsequent sections, I will draw ideas and examples from empirical studies (e.g., Bucciarelli 1994; Davis 1998) in which the attitudes and activities of engineers were observed and analyzed. I will also include some examples/quotations taken from a series of eight in-depth, semi-structured interviews of practicing engineers conducted by a handful of engineering students as part of an assigned project in my *Social and Ethical Issues in Engineering* course at Baylor University in the spring of 2014.[1] Empirical studies like this that shed light on engineering values can take forms such as interviews of engineers, observation of engineers by researchers embedded in engineering activities, or textual analysis of engineering literature, among others. Taxonomies such as that proposed here, as well as those of Hansson, van de Poel, or others, can provide a framework to help guide the collection, coding, and interpretation of data from empirical studies.

14.2 Choice of Terms: Micro, Meso, Macro, and Meta

Micro, meso, macro, and meta are hierarchical classification terms that are obviously oft-used in a wide array of contexts. My choice to employ these terms here is motivated, at least in part, from ideas drawn from two disparate sources: the literature on organizational leadership and the literature on evolutionary economics. The ways in which the terms are used in these two areas are quite distinct, and neither usage maps directly onto my proposed meanings for these terms as they relate to engineering values. Nonetheless, the ways in which the terms are employed in each of these two areas have influenced my thinking about classifying engineering values, and therefore I think it is worth providing a brief overview of these sources.

Within the literature on organizational leadership we can find the idea of micro, macro, and meta levels of leadership (Nicholls 1988). Micro-leadership refers to leadership at the level of directing particular people to accomplish specific tasks. Macro-leadership refers to leadership at the level of developing and directing organizations more generally, such that an organization might successfully pursue a variety of goals and accomplish a multitude of tasks concurrently. Both of these types of leadership typically involve the exercise of formal authority and responsibility. Meta-leadership, by contrast, is leadership that transcends any formal

---

[1] The interviews were conducted by undergraduate engineering students who received training in interview protocols. The interviewees were employed engineers who, with one exception, worked in Texas. They comprised both men and women, represented several engineering disciplines, and represented an assortment of types of industries or held governmental positions.

authority and responsibility. It is leadership that can inspire the pursuit of overarching ideals that transcend specific roles or even particular organizations.

Parallels can be drawn between these ideas about leadership and my use these terms for classifying engineering values. By micro-level values, I mean those values that are prominent in influencing specific tasks involved at the most detailed levels of engineering design, or at the level of what Louis Bucciarelli (1994) calls the *object world*: "…the domain of thought, action, and artifact within which participants in engineering design…move and live when working on any specific aspect, instrumental part, subsystem, or subfunction of the whole." By macro-values, I mean values that figure prominently at higher levels within an engineering organization, where engineers intersect more routinely with non-engineers and non-technical aspects of engineering projects. Finally, I take meta-level values to be those values that transcend particular engineering organizations and perhaps permeate the profession/activity of engineering in a widespread fashion, whether tacitly or expressly.

I have drawn no parallel in the organizational leadership literature to what I have called the meso-level of engineering values. For that, I have taken inspiration from the literature on evolutionary economics, in which the idea of a meso-economic level has been developed as a proposed bridge between the microeconomic and macroeconomic levels (Dopfer et al. 2004). Briefly, at the microeconomic level, individual agents maintain sets of rules they use to guide their interactions with other agents. At the macroeconomic level, individual agents and rules are transcended, and the focus is rather on aggregate consequences of populations of agents and rules operating within an economic ecosystem. The meso-economic level, proposed as a bridge between the micro and macro, concerns the creation, diffusion, and adoption of new rules and rule sets by and among agents, which in turn leads to dis-equilibration, change, and eventual re-equilibration at the macro-level.

I will attempt to make an analogy, at least in an abstract sense, between this notion of meso-economics, and the idea of meso-level values in engineering. At the micro-level, as I've defined it (corresponding to Bucciarelli's object world) engineers are at work on the technical bits and pieces. At the macro-level, organizational objectives are being pursued and achieved, and, perhaps simultaneously, personal and/or social ideals are being serviced at the meta-level. I propose to define a meso-level as the level connecting the micro side of engineering work to the macro/meta side. I think it is also important to make a connection between this idea of a meso-level and the idea of the dual nature of technical artifacts as articulated by Peter Kroes (2010). This structure-function dichotomy also exists at the boundary between the micro (object) world and macro/meta worlds. In the next section, I will elaborate on each of these proposed levels of engineering values, and provide examples.

14.3 Meta Level Engineering Values

The title of this article provides an example of what I mean by meta-level value in engineering – in this case, that engineering is for the sake of the benefit of humanity. Now, the question of whether this or that particular engineering work actually benefits humanity is up for debate, as is the question of whether we can even agree on how such an outcome can be measured, but what is fairly clear is that this idea is pervasive in an overarching way throughout the engineering profession. For example, the IEEE, which is the world's largest engineering professional organization, defines its mission as, "IEEE's core purpose is to foster technological innovation and excellence for the benefit of humanity." One of our engineer interviewees said something similar: "*But to me engineering is about helping people, and as engineers as a whole we design things that make life better for everybody*" (here and throughout I will use italics within quote marks to denote excerpts from our student-conducted interviews of engineers). This idea of making life better seems to be treated as more or less axiomatic by engineers, to the point of verging on boilerplate rhetoric, and perhaps it isn't always subject to as much reflection as it should be. For example, Samuel Florman (1987) writes, "Every engineer I have ever met has been satisfied that his work contributes to the communal well-being, though admittedly, I had never given much thought to why this should be so." We could hypothesize that this is due to an uncritical conflation of technological progress with human progress – what Leo Marx (1987) called the technocratic concept of progress.

Another example of an engineering meta-level value, and one that van de Poel (2015) discusses as an example of an external value, is *safety and health*. I consider it a meta-level value because of its overarching reach; almost all codes of ethics for engineering worldwide make protecting public safety and health of paramount concern to engineers. Van de Poel makes the point that safety/health is also an external value because the need or desire for it arises outside of the practice of engineering. But he also notes that it has been internalized within the practice of engineering to the extent that it has become an implicit element of the engineer's value system (or at least the value system of a canonical engineer). As one engineer put it, "*I think there's a lot of engineering that can have a direct negative impact on people, and on their safety and their lives, and you need to make sure that what you're doing is not going to hurt people.*"

Sustainability is another external value discussed by van de Poel, and one that could also be classified as a meta-level value. I say "could" because meta-level values would typically also be final values; that is, things that are valued for their own sake or, following van de Poel, things that are at least constitutive of final values. For someone who takes sustainability to heart, it is either a final value or at least constitutive of human well-being and/or the well-being of nature, and is thus a meta-level value. However, it is possible for someone to value sustainability in a more instrumental way. Green marketing, for example, could refer to a company that develops and markets environmentally sustainable products not because the company itself particularly cares about sustainability, but rather because it perceives that others do, and hence there is a market from which a profit can be made. In such a case, sustainability might be a macro-level value for the organization, but not a meta-level value.

Engineering meta-level values, while overarching, do not have to be universal. They may vary from time to time and place to place. For example, different engineering cultures exist in different locales, each with a potentially unique engineering identity, ethos, and set of allegiances. For example, Gary Downey et al. (1987) have compared national engineering cultures in France, Germany, and Japan. They note French engineers as having a strong sense of national public service, German engineers as having a strong sense of social responsibility, and Japanese engineers as having a strong sense of company loyalty. Each of these perhaps contrasts with the United States, where engineering has developed a strong sense of being an autonomous profession. It should be no surprise that such different engineering cultures might lead to the adoption of, at least some, different overarching values, or meta-level values.

Meta-level values in engineering will often have to do with either the engineer's identity or, as Michael Davis (1997) describes in his definition of *profession*, the "moral ideals" that engineers organize themselves to serve. Individual engineers will also bring their own unique meta-level values to their work. For example, an engineer might have a passion for sustainability that motivates and influences the work she does and how she does it. Or, an engineer might be a pacifist, and thus choose not to work for defense industry organizations; or, the converse might be the case. One interviewee working for a defense contractor recalled a conversation she had with an engineer from another company who expressed distaste for defense work because it leads to someone being hurt. Her response was, "*I don't care because I'm protecting my people, and that's what we're supposed to do is protect our people.*"

14.4 Macro Level Engineering Values

Engineers, for the most part, work within organizations of one type or another, and organizations pursue particular agendas and objectives, from seeking profits to providing humanitarian aid. One engineer working for a private spaceflight company characterized his company's mission this way, "*The company was founded to progress the status of humanity at the highest level.*" In this case, the company has internalized and instantiated at the macro level the meta level value of not only benefiting humanity, but actively progressing it. Along with such an overarching company mission that connects the meta and macro levels, if a company has one, there will be a whole raft of values operative for engineers at the organizational level that have to do with the objectives, structures, processes, and people within organizations. Profit, company reputation, customer needs, employee competence, just to name a few. While many macro-level values will be shared across many organizations, the particular compendium of macro-level values that characterizes an organization will be unique to each organization. I once worked for a large aerospace company that had a large sign posted at the entrance to the company's campus that read, "Building on a culture of efficiency," a clear indication of a macro-level value espoused by that organization (and likely shared by many engineering organizations). Efficiency, in this sense, is not just a technical value that operates at the micro-level of detailed technical work, but rather is a value to be cultivated throughout the organization more broadly. Many critiques of technological culture also consider that efficiency is sometimes inappropriately elevated to the meta level, to the level of being a final value, or end in itself. This indicates that values can potentially manifest themselves in different ways at different levels.

With respect to macro-level values, consider Michael Davis' (1998) empirical study of the engineer-manager relationship within companies. He reported finding three types of companies: engineer-oriented, customer-oriented, and finance-oriented. Engineer-oriented companies were ones in which the quality of the product was elevated above other values, perhaps almost thought of as a final value. In this case it is also a largely internal value, following van de Poel's classification (2015). That is, while customers certainly might appreciate quality, Davis reports that quality was not necessarily sought as primarily a means to customer satisfaction. Rather, the achievement of quality satisfied some internal desire within the company for technical accomplishment. In fact, according to Davis some companies would rather have lost a customer than sacrifice quality. One of our interviewees said about his company, "*When we hire employees, we want to make sure that they share the same philosophies that we do about quality*." This contrasts with customer-oriented companies for which customer satisfaction was the overriding value. If a customer preferred low price, and was satisfied with the commensurate low quality, then the company was happy to oblige. As one interviewee said, "*If my client's happy, then the project is a success*." In this case, the company's product design is being driven, at least to a high degree, by macro-level values that are external (arising from the customer/client), although the underlying value of giving primacy to the customer's desires is still an internal decision. The third type of company Davis reported about was the finance-oriented company, which was motivated to keep internal values such as profit, production quantity, or speed of production front and center.

In general, we would suspect that a typical organization operates on the basis of some weighted combination of internal and external values, and must exhibit some elements of each of the engineer-, customer-, and finance-orientations, even though one may dominate. On the one hand, every organization will have someone external to whom they have to pay attention, whether that's customers, clients, investors, or some other stakeholders. On the other hand, every organization will also have some internal value system that guides its operation and serves both to motivate and constrain the organization's activities in various ways. The Swiss watch industry's crisis during the 1970s perhaps provides an example of an industry grappling with competing macro-level values. The Swiss industry's financial success declined sharply during that time period due to what might be thought of as a pathological pursuit of an engineer-orientation at the expense of customer and finance concerns:

"Swiss watch companies had become absorbed in the technology of producing watches rather than thinking comprehensively about what made a customer actually want to buy a watch" (Bottger 2010). Technical knowledge and skill with mechanical watch movements, along with the quality and precision of the product, were points of pride with the Swiss watch industry that contributed to it lagging foreign competitors in both rationalizing the production processes for high quality mechanical watches, as well as diversifying its capabilities to enter the market for newer quartz movement technologies (Donzé 2014).

While many types of values that influence engineers at the macro level derive from the structure and processes of their organizations, perhaps one of the most critical clusters of values for engineering design are those values that motivate the construction of design requirements via the process of translating abstract desirables into a more concrete functional description. As mentioned earlier, van de Poel (2013) has analyzed this process as a two-step process of converting general values into prescriptive norms, and then converting the norms into specific requirements. For example, we might value a product that we can conveniently carry and transport, which might lead to a norm such as "must be portable in a pocket", which in turn might lead to a set of requirements specifying maximum dimensions, maximum weight, types of materials, etc. Van de Poel correctly points out that deriving design requirements from more abstract values is non-deductive. There are potentially many ways the translation can go, and value judgments are involved both in the translation process itself (which is an inherently interpretive process), and in choosing between competing translations. This process is complicated by ambiguity or uncertainty in the desirable attributes of the artifact or system. As one engineer put it, "*A lot of times the client really, frequently doesn't know precisely what it is they really want.*" This leads to an iterative process in which engineers much probe the client for clarification. The care and quality, or lack thereof, with which this iteration is done can significantly impact the match between the imagined and realized artifact or system.

The macro-level values generally pervading the organization can certainly influence the translation. For example, if a customer expresses the same desired product attributes to an engineer-oriented company, a customer-oriented company, and a finance-oriented company, we could imagine three quite different sets of resulting specifications. The customer oriented company would perhaps strive the hardest to assure the customer-valued attributes are translated into concrete requirements in a way that stays as faithful to the customer's perspective as possible. The engineer-oriented company is more likely to translate the customer's desires into specific requirements that, while certainly meeting the customer's criteria, also uphold the company's standards while perhaps even providing a novel technical challenge. Finally, a finance-oriented company, particularly a more cynical one, may seek to derive design requirements in a way that, while superficially adequate, gives short shrift to the customer's desires while maximizing the advantage to the company. One interviewee recalled situations where contractors would inflate costs for design features that were not really necessities, saying, "*That's kind of the game they play, is 'How much can we get out of the client?*'" In fact, some engineering codes of ethics proscribe engineers in certain situations from both specifying design requirements and performing the design due to the fact that the possibility of benefiting from performing the design might introduce certain value judgments into the specification process that aren't necessarily in the best interest of the client.

I might note that different values operating at the same or at different levels can either resonate or conflict. In the case of an engineer-oriented company, we could easily imagine the organization's commitment to quality resonating with an engineer who is motivated by technical achievement. It is no wonder that Davis reports finding engineers who "feel at home" in such organizations. Similarly, we could imagine the work of a defense company resonating with an engineer who has a strong commitment to, and support for, the military, or the work of *Engineers Without Borders* resonating with an engineer who feels passionately about humanitarianism and social justice. By contrast, we could imagine an engineer who has internalized the values of safety and quality chafing at a finance-oriented company that continually seeks to cut corners, compromise quality, and reduce costs. The recent example of the Volkswagen diesel automobile emissions scandal serves to highlight the potential conflict between internal and external values at the macro level, as well as between macro- and meta-level values.

The company designed the emissions control system of some diesel cars to detect when emissions tests were being conducted on the vehicle and to only activate the emissions controls during such tests in order to fraudulently pass the test, while otherwise during normal operation the vehicles' emissions controls would be turned off and the vehicles would not meet emissions standards. The deceptive design was apparently implemented "after realizing there was no legal way for those engines to meet tight U.S. emissions standards 'within the required time frame and budget'" (Boston et al. 2015). In this case, a set of external values related to environmental quality, made operative via translation into specifications for tailpipe gas emission levels, was supplanted by another set of values, presumably values internal to the organization related to profit, or greed, or perhaps fear of failure or embarrassment. It remains to be seen how widespread complicity in the deception was within the company, but it is clear that at least some engineers were involved since it required intentional and directed design activity to meet what amounted to a set of "shadow specifications" that was substituted for the ostensible specifications. In this case there was also a clear conflict between the macro-level values expressed in working to satisfy these "shadow specifications," and meta-level values generally associated with engineering ethics.

Another group of engineering values operative at the macro level are those associated with engineering competencies. Competencies can be categorized, among other ways, according to discipline/subject matter or according to type of task. Engineers will often have one or more subject matter areas of core competency, and perhaps other areas of peripheral competency. One engineer described the value of competency in the following way: "*Really, really, really good in your area, and a base knowledge in other engineering disciplines that you have to deal with, so you can talk to them.*" Task competencies include analytical skills, graphical visualization skills, computer skills, hands-on skills, communication skills, management skills, and the like, and are highly valued in engineering work, which obviously is why such competencies are a main focus of engineering education. While many competencies ultimately find their expression at the micro level of detailed engineering tasks – i.e., in the object world – I assign them here to the macro level because competency values are critical as the basis for judgments at the organizational level. Competencies serve in the most basic way as criteria for entry into engineering work. Thereafter, they serve as a sorting mechanism for personnel between and within organizations. While it may be hoped that all engineers possess some level of competency of all the types considered important in engineering, the reality is that different engineers will typically exhibit affinities or natural talents in particular areas. As an engineering design teacher I see this all this time with student design teams. A division of labor will naturally emerge on teams whereby each student fills a niche according to her or his particular interests and talents. The "hand-on person" will take the lead in manufacture, construction, testing. The "graphical person" will take the lead in 3-D solid modeling and engineering drawings. The "analytical person" will take the lead in making calculations or running simulations. And so forth. Similarly, organizations hire based on trying to match such interests, skills, and experience with organizational needs, and subsequently will distribute general roles and specific tasks within the organization based on the same. As discussed earlier, there is a meta-level set of values that serve to unite engineers with a global engineering identity, and in an abstract sense the set of all competencies considered important for engineers exists as a meta-level value, as does the idea of competency itself. However, the non-uniform distribution of the various competencies among engineers makes those competencies macro-level values in a concrete sense. Competencies serve to differentiate engineers, allowing each to develop a particular identity, complete with a unique repertoire of capabilities that will make him or her valuable in particular ways to particular organizations.

14.5 Micro Level Engineering Values

Although the meso level is logically next in the hierarchy, I will first address the micro level, as I think that will make it easier to subsequently discuss the meso level. As previously stated, the micro level corresponds the level of detailed engineering tasks, tasks that are carried out in Bucciarelli's object world. While engineering competency values figure prominently in organizational decision-making at the macro level, as discussed in the preceding section, their functional expression occurs at the micro level where engineers draw, compute, analyze, construct, test, and so forth. As these tasks are carried out, a host of other values come into play. It's important to have *good* data, *accurate* models, and *precise* equipment. All manner of such valuations are made about the information, tools, materials, techniques, and processes that engineers employ to carry out their work at the micro level. Van de Poel (2015) provides a list of some internal values for engineering, such as effectiveness, efficiency, robustness, reliability, and so forth. He discusses these in the context of describing qualities that the products that engineers design should have. But they also apply to the resources engineers bring to bear on the design activity itself. Engineers value having reliable test equipment, robust models that can account for many factors, and efficient protocols for carrying out the design process. Quantification is frequently a value at this level – that is, the desire to “express all variables as numbers” (Koen 2003).

Standardization is an important value at this level. Selecting parts, components, and materials is made exponentially easier knowing that there is compatibility across vendors. Standardization isn’t limited to physical artifacts, however. Electronic communications and data handling protocols, for example, are important for compatibility across electronic platforms. One engineer working in the electronics industry said, “*Technological progress is great. The problem is, it can be a free-for-all without standardization*.” Also, and this is not something unique to engineering, but is true for most professions, there is standardization of terminology and symbology. This allows engineers to exchange technical information quickly and with high fidelity. Related to standards are codes, which can be thought of as standardized, generic design requirements, often related to safety; in fact they are micro-level instantiations of the meta-level value of safety. They are generic in the sense that they apply to entire classes of designs rather than particular designs. And while some codes apply to whole products or systems, many are applied at the level of working out the details of individual components. For example, in my work in the aircraft industry I applied required margins of safety down to the level of individual rivets. While the application of codes and standards happens at the micro-level, the development of codes and standards occurs at the macro-level through either engineering professional organization activities or within and across companies.

Another micro-level instantiation of the safety meta-level value is work-checking and the need for avoiding mistakes. As Henry Petroski (1985) writes, “[I]t is the essence of modern engineering not only to be able to check one’s own work, but also to have one’s work checked and to be able to check the work of others.” Engineering organizations develop, at the macro-level, protocols for engineers to check each other’s work at the micro-level, often multiple times, in order to assure that preventable mistakes are caught. This was certainly true of my work on aircraft structural components, where multiple levels of work-checking occurred for every component. Related to checking work and eliminating errors, one value at the micro level that appeared repeatedly in our interviews with engineers, was “*attention to detail*.”

Efficiency is a value that manifests itself at multiple levels. I’ve previously mentioned the view prevalent among some critics of technological culture that efficiency has been inappropriately elevated to the status of meta-level value, or end in itself. To go along with the meta-level value idea, Walter Vincenti (1990) has written about optimization, which is an efficiency-related concept, that it is “a constant element, implicitly or explicitly, in engineering thinking. For the engineer optimization has the nature of an ethos.” One engineer interviewee had a personal, meta-level religious perspective on efficiency, saying, “*When you design a more efficient car, you bring an aspect of redemption to transportation, to culture, to society*.” I also mentioned above that one of my previous employers advertised itself as having a "culture of efficiency", which was a macro-level instantiation of the value within the organization, arguably for instrumental purposes related to the company's competitive goals more so than as a higher ideal. At the micro-level, efficiency enters in a form I have previously referred to a micro-efficiency (Newberry 2015), which is a type of efficiency that is not necessarily part of the design requirements of a product or system that is being designed, at least in any explicit way. Rather it is an intrinsic part of the engineering process at the micro level as a result of the fact that economic competition is part of the organizational process at the macro level. This is distinct from macro-efficiency, which is efficiency related to the function of the designed artifact, and is explicitly encoded in the design requirements. Design requirements can always be satisfied in a multitude of ways, many of which may be adequate to satisfy the end-user. However, there is an ineluctable pressure at the micro level of engineering work to satisfy design requirements in ways that minimize materials, manufacturing steps, labor, time, etc., not because it necessarily improves the effectiveness of the product from a use standpoint, but because it improves the competitiveness of the product from the standpoint of organizational success. After stating that "working" was the most important criteria for a design, one engineer said the next priority was to make sure it was designed "*in an elegant and efficient way so that it not only does what is needs to do*, *but that it does it well and inexpensively*."

There can also be unique personal values that engineers express at the micro level. In discussing his detailed technical work, one engineer said, "*Technically*, *it's challenging*, *which is good…I would get bored if it was easy*." As I hope this discussion reveals, the key to the way I'm defining micro-level engineering values is that they are values operative in the course of detailed-level engineering tasks, but in ways that are relatively independent of creative aspects of design. Rather they are values that underpin the generic aspects of filling in the details of a design once a design concept has been generated from the design requirements. Of course there are also many values at play in the concept generation phase of the engineering process of designing and developing. These are what I will consider in the following section on the meso level.

14.6 Meso Level Engineering Values

I've reserved discussion of the meso level till last in order to highlight its place as a bridge between the macro and micro levels. At the macro level, design requirements are constructed for the purpose of defining a functionality to be realized via the design and development of a product or system. At the micro level, engineers execute specific technical tasks such as drawing, calculating, testing. Activities at the meso level, and the values that undergird them, are what enable the micro-level tasks to result in the eventual realization of a "structural description" (Kroes 2010) of a macro-level product or system that satisfies the design requirements ("functional description"). Just as the process of translating user desired values into design requirements is non-deductive, translating a functional description of an artifact into a structural description is also non-deductive and non-unique. As one engineer said, "*For many things, just having a way to do it is pretty much good enough, but a lot of other times where it's more open it's like 'Ok, I could do it this way or I could do it that way… which one should I pick.*'" To accomplish this translation depends heavily on a few core values, and also requires applying a cadre of design heuristics that draw on a variety of other values. Here I use the term "heuristic" in the sense similar to Billy Vaughn Koen (2003): "anything that provides a plausible aid or direction in the solution of a problem but is in the final analysis unjustified, incapable of justification, and potentially fallible."

Effectiveness is most certainly a core value at the meso level. At the most fundamental level, the design process is predicated on the assumption that the desired functionality ought to be achieved. "*The most important factor is whether or not* [*your design*] *does what you said it was going to do*," according to one engineer. Some other internal values that van de Poel (2015) suggests are values relative to the object of design are quality, robustness, maintainability, and reliability. The extent to which these values are promoted is context specific. Such values may be written explicitly into the design requirements based on customer desires. They may be values internally important to the organization as a whole, such as discussed earlier with respect to engineer-oriented companies, and therefore become an implicit, unwritten part of the design requirements. They may also be values important to individual engineers working on the design and thereby become embedded in the design in an informal way. Or, they may factor into the design in some weighted combination of all three of these possibilities.

Another core value at the meso level is creativity or innovativeness – the ability to envision novel arrangements of parts and materials that will result in a desired functionality. Experience with similar designs, and understanding previous designs, is another core value. That is, there is great value in having participated in previous designs of similar products or systems, and having first hand knowledge of the materials, components, and techniques used in those designs, as well as the rationales upon which they were chosen, much of which is learned directly from more experienced colleagues. In the words of one engineer, "*The engineers that have been there five or six years are teaching the brand new ones who are coming in. And you're just constantly always teaching the new people…'Here's what you do…and here's why*'".

An important value, but one which is not necessarily obvious, is the ability to know when to stop working on the design. There is an old joke that the only way to freeze a design is to shoot the engineer. Any design can always be improved in some way, so in a sense a design is never finished. But pragmatically, to be deployed a design at some point has to be deemed "good enough." One engineer said, "*There's always improvements to be made, so you have to be careful about iterating too much because then you may just be wasting time when you already have a design that works, and you can move on to the next thing*."

Turning attention to design heuristics, one example articulated by Koen is, "Break complex problems into smaller, more manageable pieces." Sunny Auyang

(2004) expresses a similar idea in her list of engineering heuristics with "Modularize". The central value expressed in these two ways is that it is better to solve multiple small problems than one large one, so discretize the large problem. Auyang goes on to add about modularization that sub-elements should have high internal complexity and low external complexity. In other words, the inputs and outputs of the sub-element should be minimal and simple, and any complexity involved in converting input to output should be black-boxed within the element. This ensures that sub-elements are relatively independent from one another, which has the dual benefits of making them relatively easy to implement and also making them relatively interchangeable with alternative sub-elements without triggering a cascading redesign of surrounding sub-elements. Another heuristic given by Koen is, "Make the minimum decision," which is echoed again by Auyang with her rule of, "Maintain options as long as possible in the design". What these are suggesting is that during the design process one should only make those choices necessary and appropriate for the current stage of the process, but should otherwise keep alive as long as possible alternative designs, or options for alternative components, and the like, so that changes are relatively easier to make if they become necessary. Another heuristic from Auyang is, "Simplify, simplify, simplify", which I guess could be thought of as a sort of Occam's Razor for design – among competing designs, the one with the fewest elements should be selected. Another heuristic from Koen is, "make small changes in the [state of the art]", which is particularly germane to safety critical engineering products and systems. This is basically a statement of the value of engineering conservatism. Where substantial risk is involved, one's designs should not venture too far into the unknown – i.e., they should not deviate too dramatically from what's been proven to work.

While standardization and codes were discussed as important values at the micro level to guide detailed engineering work, in our interviews with engineers, standardization sometime was characterized as a negative value at the meso level. In attempting to map functional requirements into a physical morphology, standardization was sometimes referred to negatively as a constraint on innovation. "*You can get bogged down. You can get bogged down and wrapped around the axle with too many standards*, *too many requirements*," lamented one engineer.

Another example of a value that can influence design choices, and can either be explicit and external, or implicit and internal, is sustainability. The object of design could be an artifact aimed at promoting sustainability – an electric vehicle for example. In that case, the criteria for sustainability will be written explicitly into the design requirements and may originate external to the organization. Or, the object of design could be a common artifact for which sustainability is not particularly a part of its design requirements, yet the company is internally invested in producing that artifact in a way that promotes sustainability, which might then impose implicit design constraints on materials selection, manufacturing methods, and the like.

Some values can be pathological in nature when operating at the meso level. For example, the *not-invented-here-syndrome* can cause people to value internal solutions above external solutions that may be clearly superior. While the tendency to value techniques, processes, materials, or designs that are familiar, or tried-and-true, may efficiently utilize an organization's current resources and knowledge base in the short term, it can compromise the organization's ability to innovate in the long term. Pressures sometimes arise in organizations that can cause values of budget, expediency, pride, or some such, to take precedence over more legitimate design considerations, leading to design decisions that ultimately may be harmful.

14.7 Conclusion

My goal in this article has been to explore engineering values, specifically by looking at how they operate from the standpoint of a hierarchy from the more abstract and overarching to the more concrete and detailed. Toward that end, I've proposed using the familiar terminology of meta, macro, meso, and micro to describe levels within the hierarchy. These levels correspond, respectively, to the levels of the most abstract ideas about engineering identities and ideals, the values that drive engineering organizations and also those that drive the functional definitions of those artifacts which are to be designed, the values that govern how the functional descriptions of artifacts to be designed are translated into physical/structural descriptions, and the values at play in the detailed technical tasks required of engineers in the process of filling in the physical details of a proposed design.

In addition, I have tried to provide examples of values operating at the different levels taken from empirical investigations of engineers and engineering in order to highlight the importance of empirical studies as a complement to the more philosophical analysis of values. While philosophical reflection on the values that do, or should, underpin engineering, along with the ramifications of those values for the products of engineering, can prove useful to understanding, and potentially aiding, the engineering profession, an accurate picture of the values actually operating in the profession is a necessary input, and requires empirical work.

I posed a couple of specific questions in the introductory section of this essay. One question concerns the ways in which meta- and macro-level values impact meso-level values (that is, how the overarching values of engineering, along with the values expressed by engineering organizations and their stakeholders, impacts the values that inform the mapping of functional descriptions into the physical instantiations of artifacts). For example, safety is a critical meta level value for the engineering profession. That meta level value gets expressed at the macro level, whether external to the organization (a customer, client, or regulatory body specifies a particular need for safety), internal to the organization (the company has its own views about safety or risk aversion, whether for pragmatic or idealistic reasons, independent of external requirements), or both. These macro-level expressions of the value of safety will then influence the values expressed at the meso-level – that is, at the level at which functional requirements are mapped into a morphological description. For instance, the amount of conservatism in choosing a design configuration will depend on both the organization's risk aversion or tolerance as well as the requirements of customers or codes. The second question concerns the ways in which meso-level values and micro-level values interact. For example, the meso-level value of recognizing when the design is good enough and then freezing it is in tension with the micro-level values of increasing design efficiency and the satisfaction of solving technical challenges, both of which tend toward extending the design.

I also find it interesting to compare and contrast the values operative at the different levels and, in particular, to see in some cases how the same base value is instantiated in different ways at different levels. For example, take the value of *efficiency*. It has been thought by some to be an overarching meta-level value that has become representative of (in a negative way to many minds) technological culture, transcending the activity of engineering. Others have argued that it is a meta-level value internal to engineering at the highest level – that its pursuit is fundamental to the engineering ethos. At the macro-level of organizations, it is often promoted as part of organizational culture for more practical reasons of organizational competitiveness. It is also often explicitly translated into design requirements based on end-user/customer desires, and is thus encoded into the functional descriptions of artifacts at the macro-level. Conversely, it can also be explicitly flouted at the level of functional description. For example, the aircraft company for which I have worked was asked by a customer to modify a luxury VIP airplane to install a granite conference table. No aircraft engineer concerned with efficiency would ever suggest installing a granite table, and it was quite a technical challenge to do so. At the meso level, where functional descriptions in the form of design requirements get translated into physical morphologies, efficiency appears in the form of materials selection, choice of manufacturing processes, and other major design decisions aimed at realizing the physical artifact in efficient ways. So, even given the highly inefficient functional requirement of having a granite table in an aircraft, a (relatively) efficient design morphology will be sought. Finally, at the micro-level of detailed design process tasks (analyzing, drawing, testing, etc.) there is a fairly constant pressure to optimize, minimize, or reduce with respect to fine details down to level of the dimensions of minor parts.

This has only been the briefest treatment, and there are likely many more specific values that could be discussed at the various levels, as well as many more connections and insights that could be had across and between levels. One area in which this proposed taxonomy of values might find use is in the empirical study of value conflicts in engineering. One example of an oft-discussed conflict is that between an engineer's professional obligations (e.g., codes of ethics) and organizational obligations (obeying directives from superiors). Understanding what specific values are involved, as well as at what levels, or across what levels, such a conflict of values exist, might aid in both understanding the origins of the conflict and identifying possible avenues of resolution.

Acknowledgements I thank the following students who participated in conducting interviews of engineers during the spring semester of 2014. At the time they were undergraduate Mechanical Engineering or Electrical and Computer Engineering majors at Baylor University, though now all have graduated: Zach Bitting, Andrew Hoeckel, Courtney Kimutis, Peyton Lundsford, and Alexa Wilde.

References

Auyang, S. (2004). *Engineering: An endless frontier*. Cambridge, MA: Harvard University Press.

Boston, W., Varnholt, H., & Sloat, S. (2015, December 10). Volkswagen blames 'chain of mistakes' for emissions scandal. *The Wall Street Journal*. http://www.wsj.com/articles/vw-shares-up-ahead-of-emissions-findings-1449740759. Accessed 10 Dec 2015.

Bottger, P. (2010, July 1). The genius of Nicolas Hayek. *Forbes*. http://www.forbes.com/2010/07/01/nicolas-hayek-swatch-swiss-leadership-managing-watch.html. Accessed 4 Jan 2016.

Bucciarelli, L. (1994). *Designing engineers*. Cambridge, MA: MIT Press.

Davis, M. (1997). Is there a profession of engineering? *Science and Engineering Ethics, 3*, 407–428.

Davis, M. (1998). *Thinking like an engineer: Studies in the ethics of a profession*. Oxford: Oxford University Press.

Donzé, P.-Y. (2014). *A business history of the swatch group: The rebirth of Swiss watchmaking and the globalization of the luxury industry*. Basingstoke: Palgrave Macmillan.

Dopfer, K., Foster, J., & Potts, J. (2004). Micro-meso-macro. *Journal of Evolutionary Economics, 14*, 263–279.

Downey, G., Lucena, J., & Mitcham, C. (1987). Engineering ethics and identity: Emerging initiatives in comparative perspective. *Science and Engineering Ethics, 13*(4), 463–487.

Florman, S. (1987). *The civilized engineer*. New York: St. Martin's Press.

Hansson, S. O. (2013). Valuation of artefacts and the normativity of technology. In M. J. de Vries, S. O. Hansson, & A. Meijers (Eds.), *Norms in technology* (pp. 103–118). Dordrecht: Springer.

Hansson, S. O., Meijers, A., & de Vries, M. J. (2013). Introduction. In M. J. de Vries, S. O. Hansson, & A. Meijers (Eds.), *Norms in technology* (pp. 1–11). Dordrecht: Springer.

Koen, B. V. (2003). *Discussion of the method: Conducting the engineer's approach to problem solving*. New York: Oxford University Press.

Kroes, P. (2010). Engineering and the dual nature of technical artefacts. *Cambridge Journal of Economics, 34*(1), 51–62.

Marx, L. (1987). Does improved technology mean progress? *Tech Rev, 90*(1), 33–41.

Newberry, B. (2015). Efficiency animals: Efficiency as an engineering value. In S. H. Christensen, C. Didier, A. Jamison, M. Meganck, C. Mitcham, & B. Newberry (Eds.), *Engineering identities, epistemologies and values: Engineering education and practice in context* (Vol. II, pp. 199–214). Dordrecht: Springer.

Nicholls, J. (1988). Leadership in organisations: Meta, macro and micro. *European Management Journal, 6*(1), 16–25.

Petroski, H. (1985). *To engineer is human: The role of failure in successful design*. London: Macmillan.

van de Poel, I. (2013). Translating values into design requirements. In D. P. Michelfelder, N. McCarthy, & D. E. Goldberg (Eds.), *Philosophy and engineering: Reflections on practice, principles and process* (pp. 253–266). Dordrecht: Springer.

van de Poel, I. (2015). Values in engineering and technology. In W. J. Gonzalez (Ed.), *New perspectives on technology, values, and ethics: Theoretical and practical* (Boston studies in the philosophy and history of science, Vol. 315, pp. 29–46). Dordrecht: Springer.

Vincenti, W. G. (1990). *What engineers know and how they know it*. Baltimore: Johns Hopkins Press.

Weinberg, A. M. (1970). The axiology of science: The urgent question of scientific priorities has helped to promote a growing concern with value in science. *American Scientist, 58*(6), 612–617.

Byron Newberry B.S. in Aerospace Engineering, University of Alabama. M.S. in Aerospace Engineering and Ph.D. in Engineering Mechanics, both from Iowa State University. Professor of Mechanical Engineering, Baylor University. Baylor Fellow for teaching. Teaching areas include engineering design, engineering mechanics, engineering ethics, and sustainable engineering. Research interests include engineering design, engineering ethics, and philosophy of engineering and technology. Aircraft structural engineering consultant. Executive board member, National Institute for Engineering Ethics. Editor of the Springer Philosophy of Engineering and Technology book series.

Chapter 15
An Engineering Turn in Conceptual Analysis

Pieter E. Vermaas

Abstract In this chapter I discuss the notion of *technical function* as it is used in engineering and review the way in which this notion was conceptually analysed in the *Dual Nature of Technical Artifacts* program. I show that *technical function* is a term that is intentionally held polysemous in engineering, and argue that conceptual analysis informed by engineering practices should chart and explain this polysemy. The *Dual Nature* program aimed however at determining a single meaning of the term *technical function* and developed an approach to conceptual analysis, called *conceptual engineering*, for arriving at this single meaning on the basis of engineering practices. It is concluded that this conceptual engineering approach is ill-suited as conceptual analysis of the term *technical function* in engineering. This approach is nevertheless a useful tool in this analysis, since it can make explicit how specific meanings of polysemous engineering terms are useful to specific engineering tasks.

Keywords Technical function • Polysemy of engineering terms • Conceptual analysis • Conceptual engineering

15.1 Introduction

The call for an empirical turn by Kroes and Meijers (2000) propelled a development that enriched philosophy of technology with knowledge of engineering practices and with analyses and results informed by the study of these practices. The *Dual Nature of Technical Artifacts* program (2002) at the Delft University of Technology was one of its first implementations, to be followed by others on different topics in philosophy of technology and at other institutes around the world. The call also led to challenges related to doing philosophy informed by the study of engineering practices. In this chapter I discuss one such challenge as it surfaced in the Dual Nature program, and review the response to it in this program. This challenge

---

P.E. Vermaas (✉)
Department of Philosophy, Delft University of Technology, Delft, The Netherlands
e-mail: p.e.vermaas@tudelft.nl

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_15

concerns conceptual analysis and consists of addressing a phenomenon that may be taken as characteristic to engineering, namely that key terms in a discipline can be intentionally held polysemous.

The key engineering term involved in this challenge is the one of *technical function*. A main aim of the Dual Nature program was to analyse the relations between the structural and the intentional concepts employed in descriptions of technical artefacts. Such artefacts were taken as having a dual nature since they were "(i) designed physical structures, which realize (ii) [technical] functions, which refer to human intentionality" (Kroes and Meijers 2006, p. 2). Technical functional descriptions were assumed to include the descriptions of the structure and of the intentionality of technical artefacts. Hence, understanding the notion of *technical function* was a central element to the Dual Nature program. And in line with the empirical turn this understanding was to be achieved by conceptual analysis aimed at determining how engineers use the term *technical function*. An initial exploration made however clear that engineers attach multiple meanings to this term, and that there were no reasons to favour a specific one as the true meaning. If in the Dual Nature program conceptual analysis had been taken as the determination of the meanings by which terms are used in specific disciplines, this polysemy of *technical function* in engineering would have been articulated and have become the basis for capturing the dual nature of technical artefacts.[1] Yet, the program aimed at arriving at a single meaning, partly in analogy with the philosophical literature on the analysis of the notion of *biological function* and possibly also for having a clean single shot at the dual nature of technical artefacts. This aim was realised again in line with the empirical turn; rather than arguing for a single meaning through philosophical intuitions and battles about exotic cases, the analysis proceeded by considering engineering practices more broadly and deriving from these practices desiderata that this single meaning of the term *technical function* should (ideally) meet. This approach was called *conceptual engineering* to contrast it to other approaches in conceptual analysis (Houkes and Vermaas 2010a, pp. 4–5).

In this chapter I discuss the notion of *technical function* as it is used in engineering and review the conceptual engineering approach as developed in the Dual Nature program for analysing this notion. Conceptual engineering may be taken as a successful empirical turn approach to conceptual analysis if one indeed assumes that conceptual analysis should lead to single meanings for key terms. Application of the conceptual engineering approach resulted in the *ICE function theory* that identifies on the basis of detailed analyses of engineering practices a single meaning for the term *technical function* (Vermaas and Houkes 2006; Houkes and Vermaas 2010a). Conceptual engineering can however also be taken as an approach that is at odds with the empirical turn. The initial exploration of engineering practices revealed, as said, that engineers attach multiple meanings to the term *technical function*, and further analysis demonstrated that engineers for good reasons use these meanings side by side and intermittently (Vermaas 2013). Empirical findings thus gave evidence for understanding the term *technical function* as one that is intentionally

---

[1] In (Vermaas et al. 2013) the dual nature of technical artefacts is revisited taking into account the polysemy of the term *technical function*.

held polysemous in engineering. The assumption that conceptual analysis should lead to single meanings for key terms should therefore have been rejected, making conceptual engineering an ill-suited approach to the task it was developed for. I argue in this chapter that conceptual engineering nevertheless is a useful tool in conceptual analysis. If it is characteristic to engineering that it can hold its key terms intentionally polysemous, engineering poses to conceptual analysis the challenge to understand this polysemy. Conceptual engineering can be a tool for taking up this challenge by making explicit how specific meanings of key terms are useful to specific engineering tasks.

15.2 Technical Functions in Engineering

A source for analysing how engineers use the term *technical function* is the literature on engineering design methodology. This literature describes models and methods that enable or support engineers to design, and in agreement with the assumptions made in the Dual Nature program, technical functional descriptions can in such models and methods relate structural and intentional concepts in descriptions of technical artefacts. In many of these models and methods the design process starts with an intentional description of a technical artefact consisting of a specification of the goals a client wants to realise with the technical artefacts to be designed, often accompanied with further technical and physical requirements the artefact has to meet. The designing engineers then translate these goals and requirements into descriptions of the technical functions of the artefact to be designed, and decompose these technical functions in subfunctions. And finally the engineers give a structural description of the artefact, which determines the physical composition and geometric dimensions the artefact and its components should have for, together, being able to executing these functions and subfunctions. This latter description is handed over to manufacturing for actually producing the technical artefact.

Yet, in spite of the consensus about the role technical functional descriptions have, these models and methods may attach different meanings to the term *technical function*. In the regularly discussed *Function-Behavior-Structure* model of design (Gero 1990) functions are defined as “the design intentions or purposes” (Gero et al. 1992). And in the well-developed *Functional Basis* account for modelling technical functions and functional decomposition a product function is taken as “the operation the product performs on incoming flows of materials, energy and information” (Stone and Wood 2000). By that second account technical functions may be seen as the intended physical behaviour of technical artefacts, making clear that according to the engineering literature there is substantial freedom in the meaning engineers attach to the term *technical function*.[2] Moreover, some design methodologists explicitly put forward that this freedom is part of engineering. In, for instance

---

[2] A review of the modelling of technical functions in engineering even listed 18 different approaches, although it should be noted that in some of the approaches a similar meaning may be attached to the term *technical function* (Erden et al. 2008).

(Chandrasekaran and Josephson 2000), it is analysed that the meaning attached to the term *technical function* can vary over a spectrum consisting of "the intended behaviour" of a technical artefact, "the intended effect on its environment" the artefact has, or any combination of the two. And in (Srinivasan and Chakrabarti 2009) a design method is presented in which three different meanings of *technical function* are discerned, which all have their use in the method.

One may argue that the engineering literature on design models and methods does not give direct access to actual engineering practices since design methodology is normative and thus merely prescribing how engineers could or should design. This literature can nevertheless be taken as indicative to how the term *technical function* is actually used by engineers. Models and methods for designing are often abstracted from actual practices of (good) engineering design, and also reveal how engineers will design in the future when some of these models and methods are adopted in engineering. Hence, the analysis of design models and methods provides empirical evidence that in engineering *technical function* is not be taken as a univocal term. One may consider taking the notion of *technical function* as a family resemblance concept. Yet this notion does not precisely fit Wittgenstein's (1953, Sect. 66) original characterisation of such concepts. The lack of a single meaning is not due to a vagueness of the meaning of the term *technical function*; in design methodology the term has rather a set of well-defined meanings. A more apt conclusion is that *technical function* is a polysemous term.

This result may be taken as giving support to the usefulness of conceptual analysis in engineering construed as disambiguating meanings attached to terms and arriving at single meanings for these terms: *technical function* is a key term in engineering, hence arriving at a single precise meaning seems a valuable contribution to a clear and well-defined conceptual basis for engineering. Conceptual analysis in this sense would furthermore not only serve a wish for clarity in science or philosophy; also in the engineering literature one can find pleas for such "consolidation" of key engineering terminology (e.g., Birkhofer 2011). The result however poses a methodological problem to this conceptual analysis, since engineering does not give a basis for advancing one true meaning of *technical function*. One can disambiguate these meanings but the separation of different meanings seems not to be accompanied with reasons for favouring one over the others. A way out of this methodological problem could be to find grounds outside of engineering for arguing that one specific engineering meaning of the term *technical function* is the true one for engineering. One could, for instance, choose a starting point in existing conceptual analyses in philosophy of the more general term *function*, and find in those analyses reasons to embrace a specific engineering meaning of *technical function*. In philosophy the term *biological function* has been a topic of extensive analysis which has led to accounts that advance single meanings (sometimes two) for this term. Some of these accounts have been generalised and now also advance particular meanings for the term *technical function* (e.g., Dennett 1971, 1990; Millikan 1984, 1993; Neander 1991a, b; Preston 1998, and more recently Krohs 2009; Longy 2012). Hence, by subscribing to one of these accounts, one fixes also the meaning of *technical function*. Yet, proceeding in this way would mean taking distance from the empirical turn in the philosophy of technology. This way out would mean understanding technical functions not just on the basis of analyses of engineering practices but also on philosophical battles about accounts of biological functions and views about how the realms of biology and technology are related. Conceptual analysis may certainly in addition to merely describing the meanings attached to terms, also involve a more normative element of adjusting those meanings. Yet, using philosophical analyses of biology for removing the polysemy of key terms in engineering seems at odds with the empirical turn.

15.3 Conceptual Engineering

The response in the Dual Nature program to the methodological problem that conceptual analysis may lead to fixing a meaning of the term *technical function* on the basis of grounds other than engineering practices, has been to embed the analysis in a larger project of giving a theory of technical artefacts. It was explored what such a theory should be able to capture and this exploration led to the formulation of four desiderata that functional descriptions of technical artefacts should ideally meet. And by deriving these desiderata from the ways engineers describe technical artefacts the desiderata originated in engineering practices, restoring compliance with the empirical turn. This response was called *conceptual engineering* (Houkes and Vermaas 2010a, pp. 4–5) for the desiderata can be taken as specifications that constrain the meaning of the term *technical function*, in analogy to the specifications that in engineering design constrain the technical artefacts to be designed. Conceptual engineering is in principle just a form of conceptual analysis, but in contrast to other forms, it does not immediately focus on definitions of the term at hand or on usages from which those meanings can be derived; conceptual engineering rather advances by first determining what kinds of constraints, or desiderata, those meanings should meet.

Conceptual engineering of the term *technical function* can be taken as an approach in conceptual analysis that leads to clear empirically informed desiderata that fix one specific meaning of *technical function*. But it is also an approach that allows for some freedom in fixing this meaning since it is in principle possible to derive from engineering practices other desiderata that fix a different meaning of *technical function*. Hence, conceptual engineering can be taken as an approach that makes through the chosen desiderata the empirical basis explicit for favouring a specific meaning of *technical function* from the multitude of meanings that engineers attach to this term. The desiderata that were chosen in Houkes and Vermaas (2010a, pp. 5–8) are given in Table 15.1.

The account of technical functions that was developed using these four desiderata is the *ICE function theory*, part of an action-theoretical description of (the using and the designing of) technical artefacts (Houkes and Vermaas 2010a). The central concept in this action-theoretical description is that of *use plan*. Following standard action theory a plan is taken as a goal-directed series of considered actions, and a use plan of an object *x* as a plan where some of the actions involve interacting with the object. The *use* of an object *x* by an agent can then be described as the carrying

Table 15.1 Four desiderata for a theory of artefacts and their technical functions

The proper-accidental desideratum:
A theory of artefacts should allow that artefacts have a limited number of enduring proper [technical] functions as well as more transient accidental functions
The malfunctioning desideratum:
A theory of artefacts should introduce a concept of a proper [technical] function that allows malfunctioning
The support desideratum:
A theory of artefacts should require that there exists a measure of support for ascribing a [technical] function to an artefact, even if the artefact is dysfunctional or if it has a [technical] function only transiently
The innovation desideratum:
A theory of artefacts should be able to ascribe intuitively correct [technical] functions to innovative artefacts

Houkes and Vermaas (2010a, p. 5)

Table 15.2 Central definition of technical function in the ICE theory

An agent a ascribes the capacity to ϕ as a [technical] function to a [technical] artefact x, relative to a use plan p for x and relative to an account A, iff:
I. The agent a has the capacity belief that x has the capacity to ϕ, when manipulated in the execution of p, and the agent a has the contribution belief that if this execution of p leads successfully to its goals, this success is due, in part, to x's capacity to ϕ;
C. The agent a can justify these two beliefs on the basis of A; and
E. The agents d who developed p have intentionally selected x for the capacity to ϕ and have intentionally communicated p to other agents u.

In the version of Vermaas and Houkes (2006)

out of a use plan for that object x. And the description of *design* can then be split in two: designing can be described as *plan design* consisting of developing a use plan for an object x for realising a specific goal, and as *product design* consisting of giving the physical description of the object x. Finally a distinction can be made between *proper use* of an object x and its *improper use*. Proper use is use in accordance with a socially accepted use plan, say because the plan is developed by professional designers, or because it is a socially well-entrenched plan. Improper use is use by a use plan that does not have this acceptance, say because it is an improvised plan made up by an agent for realising an idiosyncratic goal.

Based on this action-theoretical description of using and designing, a technical function of a technical artefact x can be roughly described as a role the artefact x plays in a use plan for the artefact that is justified and communicated to prospective users. Spelled out in detail a technical function is a specific capacity to ϕ that an agent ascribed in a justified manner to the technical artefact x if three conditions are satisfied.[3] The central definition of the ICE function theory is given in Table 15.2.

---

[3] These conditions are labelled I, C and E and refer to abstracted versions of three general theories of functions available in the philosophical literature: the I(ntentionalist), C(ausal-role), and E(volutionist) theories.

In Houkes and Vermaas (2010a, Sect. 4.3 and Chap. 5) it is argued to what extent this ICE function theory meets the four desiderata for a theory of technical artefacts as given in Table 15.1.[4] In short, the *proper-accidental desideratum* is met by defining proper technical functions as technical functions that an agent ascribes to technical artefacts relative to use plans that capture proper use of the artefacts, and by defining improper technical functions as technical functions an agent ascribes relative to use plans that capture improper use. The *malfunctioning desideratum* is met, for instance, because under specific circumstances an agent may reasonably believe and justify that a technical artefact has a certain capacity to ϕ that contributes to realising the goal of its use plan (conditions I and C) for which it is designed and presented (condition E) although in actuality the artefact does not have that capacity. The *support desideratum* is met through the conditions I and C. And, finally, the *innovation desideratum* is met because a technical artefact that is (successfully) designed for realising a use plan by a novel capacity to ϕ, can after communication of this use plan to other agents be ascribed this capacity as a technical function.

Before assessing this ICE function theory further, a brief summary of how it arrives at a specific meaning of *technical function*. In engineering, as describe above, the term *technical function* has various meanings, including design purposes and intended physical behaviour. With conceptual analysis one can differentiate and articulate these meanings, yet not favour one as the right meaning. The approach of conceptual engineering used for arguing for the ICE function theory, overcomes this deadlock by drawing from engineering practices a set of desiderata that impose constraints on the meaning of *technical function*. The ICE function theory meets these desiderata and defines technical functions as capacities of technical artefacts, which corresponds more or less to the meaning "the intended effect on its environment [of the artefacts]" as discerned in Chandrasekaran and Josephson (2000). And the support to favour this particular meaning as the right one for *technical function* is that the ICE function theory meets the desiderata. As said, this method of conceptual engineering allows that other desiderata could be chosen, which opens up the possibility to argue for an alternative account of technical functions that may favour other engineering meanings.[5]

---

[4] In (Houkes and Vermaas 2014) the ICE function theory and the action-theoretical description of using and designing is extended to an analysis of the concept of technical artefact itself.

[5] A complication can be that an alternative account that favours such a different meaning for *technical function* may also meet the four desiderata given in Table 15.1. Let us for the sake of argument ignore this latter possibility, and accept that the ICE theory is the only account of technical functions that meets the four desiderata of Table 15.1, such that these desiderata are indeed giving sufficient support to the ICE theory and to the position that the meaning it attaches to the term *technical function* is the right one.

15.4 Reception of the ICE Function Theory

The ICE function theory was framed as an engineering account of technical functions, and was as such presented in philosophy and engineering. The reception of this theory in these two disciplines was however markedly different.

In philosophy the ICE function theory was received in a rather standard way. It was taken as a possible account of technical functions that defines a specific meaning of this term based on studies of engineering practices. And as such it became subject of standard philosophical procedure: the ICE function theory was applied to address philosophical problems, it was criticised and it was compared with other accounts of technical functions.

Within the Dual Nature program the ICE function theory was used for analysing how technical functional descriptions relate structural and intentional concepts in descriptions of technical artefacts. Using the metaphor of a bridge (beyond technical feasibility) it was argued that technical functions act as "two-sided drawbridges" between structural and intentional descriptions of technical artefacts (Vermaas and Houkes 2006). A full technical functional description includes details about how the technical functions of technical artefacts are enabled by the physical structure of the artefacts (in essence a description of how the capacities to ϕ that are ascribed as technical functions are capacities of the physical structure of the artefacts) and by including information on how technical functions contribute to the intentional goals users can realise with the artefacts (in essence a description how the capacities to ϕ ascribed as functions make the use plans of the artefacts means to obtain these goals). Yet functional descriptions can also be partial. In engineering product design technical functional descriptions may be limited to descriptions of how technical functions are enabled by the physical structure of technical artefacts; functional descriptions then *highlight* the structural description of technical artefacts and black-box their intentional description – as if the conceptual bridge is drawn up to *cloak* the intentional part of the dual nature of the technical artefact. Yet in more conceptual plan design, technical functional descriptions may be limited to descriptions of how technical artefacts can be used to realise goals; thus *highlighting* the intentional description – as if the conceptual bridge is drawn up to *cloak* the structural side of technical artefacts.

Further applications of the ICE function theory involved analysis and critiques of positions in metaphysics to characterise artefact kinds by means of their technical functions (Houkes and Vermaas 2004, 2014). The ICE function theory was moreover criticised (by, e.g., Preston 2003, 2013; Hansson 2006; Schyfter 2009; Kroes 2012, Chap. 3) and compared with and contrasted to rival accounts of technical functions in philosophy, specifically the earlier mentioned accounts of biological and generalised functions (e.g., Houkes and Vermaas 2010b; Vermaas and Houkes 2013).

In contrast to this more regular uptake of the ICE function theory in philosophy its reception in engineering was rather indifferent and minimal. In spite of the assumed usefulness of a more precise characterisation of the term *technical function* to a well-defined conceptual basis for engineering, the ICE function theory hardly has been discussed in detail in the engineering literature; the few responses it did elicit was that it was a rather complex account for practical application in engineering practice.

A somewhat self-affirming philosophical response to this minimal reception could be that engineering is not yet ready for conceptual precision, in spite of the pleas within engineering for consolidating key terminology. A second more self-critical response is to conclude that the project of conceptual analysis of the term *technical function* was simply ill-judged. The Dual Nature effort included the assumption that key terms should have a single precise meaning. Yet, the study of engineering practices gave evidence that exactly that assumption did not hold for the term *technical function*; the literature on engineering design, specifically (Chandrasekaran and Josephson 2000; Srinivasan and Chakrabarti 2009), clearly demonstrated that this term is deliberately assigned different meanings in engineering. Hence, when the empirical turn was taken at heart, it could have been anticipated that a philosophical account of technical functions that advances a single meaning could at best provide a conceptual analysis of *one* of the different meanings of *technical function*. On this second response to the minimal reception of the ICE function theory in engineering, the challenge for philosophy is to understand why engineers are keeping a key term deliberately polysemous.

15.5 Technical Functions in Engineering Revisited

Later research after the Dual Nature program focussed on understanding why in engineering the term *technical function* is used with different meanings. Generalising the two-sided drawbridge-analysis of how technical functional descriptions can relate and cloak the structural and intentional descriptions of technical artefacts, it was argued that engineers can adjust the meaning of technical functions depending on whether and how in design they simplify their reasoning (Vermaas 2009). When no simplification is adopted, the reasoning in design is extensive by including various different concepts: engineers describe the *goals* clients want to realise with the technical artefacts, the *actions* users have to carry out with these artefacts for obtaining these goals, the *capacities* the artefacts have to have for making these actions effective, the *physical behaviour* the artefacts have to display for having these capacities, and finally the *structural properties* of these artefacts for enabling them to display this behaviour. In such extensive reasoning schemes it makes sense to let the term *technical function* refer to the capacities of artefacts, as in the ICE function theory, turning it in a concept distinct to those of goals, physical behaviour and structural property, and making functional descriptions indeed descriptions that relate the intentional goals to be realised by technical artefacts with their structural behaviour and properties. Yet, in engineering design such extensive reasoning is often avoided, for instance, to make this reasoning more economical and efficient. In the afore mentioned Function-Behavior-Structure model of design (Gero 1990) actions with technical artefacts are not considered, and the description of the goals of clients and the technical functions of artefacts are effectively equated, turning design reasoning, as the model's name already suggests, into primarily reasoning about the functions, physical behaviour and structural properties of artefacts. In this simplification references to actions are thus cloaked in design reasoning and technical functions can refer to the goals of clients; in the Function-Behavior-Structure design model technical functions are taken as "the design intentions or purposes." Another simplification that can be encountered in engineering design is reasoning in which both the actions with technical artefacts and the physical behaviour of these artefacts are cloaked, as in, e.g., the model of design given in Stone and Wood (2000). The conceptual distinction between the capacities of technical artefacts and their physical behaviour is then absent, allowing that technical functions refer much more to the physical behaviour of technical artefacts; function is by Stone and Wood (2000) taken as "the operation the product performs on incoming flows of materials, energy and information."

Technical functional descriptions can thus be employed in engineering in various design methods and the meaning of *technical function* can be flexibly adjusted to the simplifications advanced in these methods. Generalising this analysis, it was argued that the meaning of the term *technical function* is in engineering adjusted to the task for which technical functional descriptions are used (Vermaas 2013). Not only specific design methods require tweaking the precise meaning of this term, but other tasks involving technical functional descriptions, as, e.g., archiving engineering knowledge about technical artefacts, may also be supported by adopting a specific meaning of *technical function*. This practice in engineering is however possible only if engineers indeed have different meanings available for the term *technical function*.

In a special journal issue aimed at reviewing and assessing the existence and use of the different meanings of *technical function* in engineering (Vermaas and Eckert 2013) further evidence of this flexible use of technical functions can be found. In Eisenbart et al. (2013) is it described how in different engineering subdisciplines different meanings are used. This implies that in interdisciplinary collaborations in design, engineers can attach different meanings to shared technical functional descriptions. This phenomenon that engineers for different subdisciplines interpret descriptions differently has already been described within the engineering literature, and is analysed as supporting creativity in design (Bucciarelli 1994). A picture that emerges is that in engineering design it is beneficial to let slide the precise meaning of the term *technical function*. In the initial phase of conceptual design a too early adoption of a specific structural solution direction for the technical artefact to be designed is avoided by using a meaning for *technical function* that refers primarily to the goals the artefact has to realise. In later phases, when the structure of the technical artefact is to be fixed and detailed, technical functions can start to refer to the capacities the artefact has to have, and then in the end to its physical behaviour. The meaning of the term *technical function* is thus flexibly adjusted to the task at hand during design.

15.6 Conceptual Engineering Revisited

Conceptual engineering or, more generally, conceptual analysis approaches in which the assumption is made that key terms should end up with single meanings, seem to be ill-suited for understanding the notion of *technical function* in engineering in a way faithful to the empirical turn. By conceptual engineering one may be able to analyse this notion on the basis of studies of engineering practices, yet, as demonstrated by the ICE function theory, by assumption it leads to advancing a single meaning for the term, whereas studies of engineering practices show that *technical function* has various meanings.

Attempts at understanding technical functions in engineering that do take the empirical turn serious should acknowledge this polysemy and drop the assumption that key terms have a single meaning. Conceptual analysis would then lead to fixing the possible engineering meanings of the term *technical function*, thus charting the polysemy and enabling further research on explaining this polysemy and on relating the different meanings to each other (as attempted in, e.g., Carrara et al. 2011).

In this project the approach of conceptual engineering seems not to have an added value to conceptual analysis. An argument to nevertheless retain this approach is that conceptual engineering can be used to determine those meanings that suit a specific engineering task. If the specific meaning by which engineers use the term *technical function* indeed depends on the task for which they employ this term, a full understanding of technical functions includes explaining how specific tasks favour specific meanings. With conceptual engineering this can be done by including this use for a task as a separate "*task x desideratum*." Conceptual engineering then amounts to finding those engineering meanings of *technical function* that fit this task desideratum and other desiderata, possibly similar to the ones given in Table 15.1. In principle one can also by regular conceptual analysis label the different engineering meanings by the tasks for which they are used. Doing it with conceptual engineering has the added value of making this analysis more explicit and normative. The *task x desiderata* make explicit the relation between the meanings of *technical function* and the tasks for which engineers use them, and the argument that a specific meaning meets a specific *task x desideratum* establishes that this meaning is also suited for carrying out the task.

In this chapter I have discussed the approach of conceptual engineering in the context of analyses of the notion of *technical function* in engineering. Conceptual engineering may also have its value for analysing other key terms in engineering. The term *technical artefact* itself may be another candidate, since also that term seems to resist being captured by a single meaning (e.g., Borgo et al. 2014). One can envisage that this polysemy can also be understood as due to the existence of different tasks for which descriptions of entities as technical artefacts are employed. A useful demarcation of the natural and technological realm by designating specific entities as technical artefacts, may be different for engineers working in more traditional disciplines like civil and mechanical engineering, as compared to those working in novel disciplines like informatics, systems engineering, synthetic engineering and human enhancement. Conceptual engineering may be the approach to make this task dependency of meanings of engineering key terms explicit on the basis of analyses of engineering practices.

Conceptual engineering may also serve this role outside of engineering. In biology conceptual analysis of the term *biological function* also revealed considerable polysemy (Wouters 2003) despite insistence in the philosophy of biology to arrive at a single more monolithic analysis of this term. Here again conceptual engineering may be the approach to understand this polysemy in relation to the different (explanatory) task for which biological functional descriptions are used by biologists. Finally, returning briefly to the ICE function theory: if this account is seen in its context of the Dual Nature program, it can be argued that the meaning it identified for *technical function* was also related to a task, namely the philosophical task of understanding the dual nature of technical artefacts. Hence, with hindsight, a fifth desideratum may be added to the four given in Table 15.1, stating something in the direction that the concept of *technical function* should allow connecting the structural and intentional descriptions of technical artefacts. Such a fifth desideratum would make explicit what the aim of the ICE function theory is, and explain the asymmetric reception of this theory in philosophy and engineering.

The impact of Kroes and Meijers' call for an empirical turn may thus extend beyond the philosophy of technology it was aiming at. The call led to exploring approaches for addressing the challenge to understand key terms that are intentionally held polysemous, and the conceptual engineering approach it led to may be of use in general, leading to an engineering turn in conceptual analysis.

References

Birkhofer, H. (2011). Conclusions. In H. Birkhofer (Ed.), *The future of design methodology* (pp. 291–295). London: Springer.

Borgo, S., Franssen, M., Garbacz, P., Kitamura, Y., Mizoguchi, R., & Vermaas, P. E. (2014). Technical artifacts: An integrated perspective. *Applied Ontology, 9*, 217–235.

Bucciarelli, L. L. (1994). *Designing engineers*. Cambridge, MA: MIT Press.

Carrara, M., Garbacz, P., & Vermaas, P. E. (2011). If engineering function is a family resemblance concept: Assessing three formalization strategies. *Applied Ontology, 6*, 141–163.

Chandrasekaran, B., & Josephson, J. R. (2000). Function in device representation. *Engineering with Computers, 16*, 162–177.

Dennett, D. C. (1971). Intentional systems. *Journal of Philosophy, 68*, 87–106.

Dennett, D. C. (1990). The interpretation of texts, people and other artifacts. *Philosophy and Phenomenological Research, 50*(S), 177–194.

Eisenbart, B., Gericke, K., & Blessing, L. (2013). An analysis of functional modeling approaches across disciplines. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 27*, 281–289.

Erden, M. S., Komoto, H., Van Beek, T. J., D'Amelio, V., Echavarria, E., & Tomiyama, T. (2008). A review of function modelling: Approaches and applications. *Artificial Intelligence for Engineering Design, Analysis, and Manufacturing, 22*, 147–169.

Gero, J. S. (1990). Design prototypes: A knowledge representation schema for design. *AI Magazine, 11*(4), 26–36.

Gero, J. S., Tham, K. W., & Lee, H. S. (1992). Behaviour: A link between function and structure in design. In D. C. Brown, M. B. Waldron, & H. Yoshikawa (Eds.), *Intelligent computer aided design* (pp. 193–225). Amsterdam: Elsevier.

Hansson, S. O. (2006). Defining technical function. *Studies in History and Philosophy of Science, 37*, 19–22.

Houkes, W., & Vermaas, P. E. (2004). Actions versus functions: A plea for an alternative metaphysics of artifacts. *The Monist, 87*, 52–71.

Houkes, W., & Vermaas, P. E. (2010a). *Technical functions: On the use and design of artefacts*. Dordrecht: Springer.

Houkes, W., & Vermaas, P. E. (2010b). Théories des fonctions techniques: combinaisons sophistiquées de trois archétypes. In J. Gayon & A. de Ricqlès (Eds.), *Les Fonctions: Des Organismes aux Artefacts* (pp. 381–403). Paris: Presses Universitaires de France.

Houkes, W., & Vermaas, P. E. (2014). On what is made: Instruments, products and natural kinds of artefacts. In M. Franssen, P. Kroes, T. A. C. Reydon, & P. E. Vermaas (Eds.), *Artefact kinds: Ontology and the human-made world* (pp. 167–190). Dordrecht: Springer.

Kroes, P. (2012). *Technical artefacts: Creations of mind and matter – a philosophy of engineering design*. Dordrecht: Springer.

Kroes, P., & Meijers, A. (2000). Introduction: A discipline in search of its identity. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (pp. xvii–xxxv). Amsterdam: JAI/Elsevier Science.

Kroes, P., & Meijers, A. (2002). The dual nature of technical artifacts: Presentation of a new research program. *Techné, 6*(2), 4–8.

Kroes, P., & Meijers, A. (2006). The dual nature of technical artefacts. *Studies in History and Philosophy of Science, 37*, 1–4.

Krohs, U. (2009). Functions as based on a concept of general design. *Synthese, 166*, 69–89.

Longy, F. (2012). Artifacts and organisms: A case for a new etiological theory of function. In P. Huneman (Ed.), *Functions: Selection and mechanisms* (pp. 185–211). Dordrecht: Springer.

Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. Cambridge, MA: MIT Press.

Millikan, R. G. (1993). *White queen psychology and other essays for Alice*. Cambridge, MA: MIT Press.

Neander, K. (1991a). Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science, 58*, 168–184.

Neander, K. (1991b). The teleological notion of "function". *Australasian Journal of Philosophy, 69*, 454–468.

Preston, B. (1998). Why is a wing like a spoon? A pluralist theory of functions. *Journal of Philosophy, 95*, 215–254.

Preston, B. (2003). Of marigold beer: A reply to Vermaas and Houkes. *British Journal for the Philosophy of Science, 54*, 601–612.

Preston, B. (2013). *A philosophy of material culture: Action, function, and mind*. New York: Routledge.

Schyfter, P. (2009). The bootstrapped artefact: A collectivist account of technological ontology, functions, and normativity. *Studies in the History and Philosophy of Science, 40*, 102–111.

Srinivasan, V., & Chakrabarti, A. (2009). SAPPhIRE: An approach to analysis and synthesis. In *eProceedings of ICED*, Palo Alto, CA, 24–27 Aug 2009 (pp. 2.417–2.428).

Stone, R. B., & Wood, K. L. (2000). Development of a functional basis for design. *Journal of Mechanical Design, 122*, 359–370.

Vermaas, P. E. (2009). The flexible meaning of function in engineering. In *eProceedings of ICED*, Palo Alto, CA, 24–27 Aug 2009 (pp. 2.113–2.124)

Vermaas, P. E. (2013). On the co-existence of engineering meanings of function: Four responses and their methodological implications. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 27*, 191–202.

Vermaas, P. E., & Eckert, C. (2013). My functional description is better! *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 27*, 187–190.

Vermaas, P. E., & Houkes, W. (2006). Technical functions: A drawbridge between the intentional and structural nature of technical artefacts. *Studies in History and Philosophy of Science, 37*, 5–18.

Vermaas, P. E., & Houkes, W. (2013). Functions as epistemic highlighters: An engineering account of technical, biological and other functions. In P. Huneman (Ed.), *Functions: Selection and mechanisms* (pp. 213–231). Dordrecht: Springer.

Vermaas, P. E., van Eck, D., & Kroes, P. (2013). The conceptual elusiveness of engineering functions: A philosophical analysis. *Philosophy and Technology, 26*, 159–185.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.

Wouters, A. G. (2003). Four notions of biological function. *Studies in History and Philosophy of Biology and Biomedical Science, 34*, 633–668.

Pieter E. Vermaas M.S. in Theoretical Physics from University of Amsterdam, and Ph.D. in Philosophy of Physics from University Utrecht. Currently Associate Professor with the Philosophy Department of Delft University of Technology, the Netherlands. Research interests: the concept of technical function; technical artefacts in engineering and metaphysics; design methodology; design for values. Editing for the *NanoEthics and Philosophy and Technology* journals. Editor-in-Chief of the Springer *Philosophy of Engineering and Technology* book series.

Chapter 16
The Concept of Function in Critical Theory of Technology

Andrew Feenberg

Abstract The concept of function is a hinge between the material world and the cultural world. Analytic philosophy of function has made considerable progress in the conceptual analysis of function, but it has not considered the link between function and culture. That is the purpose of this chapter. We know from social constructivist investigations of technologies that the problems to which technical solutions are addressed depend on the interpretations of actors with the power to influence design. Corresponding functions are designed into technical artifacts. The interpretations and therefore the functions depend on the cultural framework within which the actors understand their own needs and the constraints of the environment. The theory of function must situate it in relation to the culture and way of life it serves. Heidegger and Lukács offer perspectives on this relation. This chapter explains their approach as it has been appropriated in critical theory of technology.

Keywords Function • Technology • Heidegger • Lukács • Rationality

16.1 Introduction

What is a technical object? How is a rock changed when it is used to crack open a shell? What transformation does a branch undergo when it is swung high to knock down an out of reach piece of fruit? Clearly the objective properties of these simple objects are not altered by their technical employment. The functions they have acquired are purely relational, that is, they would not exist except for the role human beings assign the objects in their practices. But the assignment cannot be arbitrary, as it would be in the case of a purely social function such as the meaning of a new word in the language. In these technical examples, there is a relation between the assignment and the properties of the objects. Those properties are part of what

---

A. Feenberg (✉)
Simon Fraser University, Vancouver, BC, Canada
e-mail: feenberg@sfu.ca

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_16

motivates the choice of these specific types of objects. The stick only acquires its fruit-gathering function because of its weight and length, the rock its shell-opening function because of its hardness. Technical objects have a foot in two worlds, a world of human intentions and a world of objective properties.

The analytic theory of technical function has attempted to tease out the exact nature of this relation, in some cases emphasizing the objective properties, in other cases human intentions, and in the most convincing formulations, achieving a balance between the two sides of the relation. The purpose of these theories is to explain how engineers use the word function, or how the word is used in everyday speech, or in both contexts where the theorists can identify a common basis. This approach abstracts from many social and cultural aspects of function in order to achieve conceptual precision.

Wybo Houkes and Pieter Vermaas have proposed the "use-plan" or "ICE" theory of function which synthesizes many analytic contributions.[1,2] Their concept of plan is meant not as a psychological description but as a way of reconstructing the nature of artifact use after the fact. Thus the theory allows for informal and incomplete interactions with objects which are more common than the prior elaboration of a detailed plan. In their theory the subjective side of functionality consists in beliefs and intentions together constituting a use-plan, while the objective side of functionality consists in specific physical properties. The beliefs about those properties must be justified for the functional ascription to count as rational. Justifications may be based on direct experience or on information obtained from experts. Either alternative makes it possible to formulate a rational plan for using an object. The theory is tested against several desiderata, such as whether it can support a distinction between "proper use" and occasional or accidental use, and whether it can explain the malfunctioning of useful objects.

> In summary, we arrive at an analysis of artefacts as objects with a twofold dual nature: they are objects that have intentional characteristics and that have physical characteristics, as well as objects that are used and that are man-made. Functional descriptions are relevant to the first, intentional-physical duality since these descriptions allow users and engineers to connect and disconnect teleological and structural descriptions of artefacts. Hence, technical function is a useful concept, that serves as a conceptual hinge between the two natures of artefacts.[3]

Although I can agree that the assignment of a function does presuppose belief about the properties of the object, I want to better understand what we *do* when we envisage the world with a technical intention. What is the orientation of the subject toward the object in this particular kind of belief? As I will show, answering this question involves understanding the specific type of object that underlies the assignment of function and the corresponding form of subjectivity.

---

[1] Wybo Houkes and Pieter E. Vermaas (2010) *Technical Functions*: *On the Use and Design of Artefacts*. Dordrecht: Springer.

[2] Another important synthesis of a wide range of literature on function is Beth Preston (1998) Why is a wing like a spoon? a pluralist theory of function, *Journal of Philosophy*, 95(5), 215–254.

[3] Houkes and Vermaas (2010, pp. 11–12).

We know from social constructivist investigations of technologies that the problems to which technical solutions are addressed depend on the interpretations of actors with the power to influence design. Corresponding functions are designed into technical artifacts. The interpretations and therefore the functions depend on the cultural framework within which the actors understand their own needs and the constraints of the environment. The theory of function must situate it in relation to the culture and way of life it serves. This has implications for our conception of modernity as a rational form of society and for the related notion of progress. Martin Heidegger and Georg Lukács have written about technology in ways that reflect an implicit concept of functionality. In so doing they work with a very different ontology from analytic philosophy. They understand the functional object in terms that have much in common with neo-Kantianism and phenomenology. The object is not "real" in any of the usual senses of the term, but rather it is the correlate of a subjective apprehension or intention. This type of object is not simply a sum of physical properties but is what might be called a "relevance structure." Such objects are routinely invoked in discussing academic disciplines: physics has matter in motion as its object, biology, living things, and so on.

These objects are not subjective but nor are they substantial entities; they are meaningful, coherent aspects of the infinite stuff of experience. This conception of artifacts as objects does not contradict the analytic philosophers' concern with physical properties in the attribution of function, but it calls attention to the selection that privileges some types of properties over others. For philosophers of technology in the Continental tradition artifacts are objects of the subjects of such selection. In what follows I will show how the technical object and subject is construed in Heidegger, the early Marxist Lukács, and in my own attempt to synthesize their contributions in the critical theory of technology.

16.2 Technical Function and World in Early Heidegger

Heidegger developed two theories of technical artifacts, an early one based on craft and a later one that concerns modern technology. The early theory as presented in *Being and Time* is a phenomenology of the everyday technical lifeworld. By "world" Heidegger means a system of meaningful entities that refers back to an agent capable of interpreting its environment, entertaining purposes and acting. This phenomenological concept of world is difficult to separate from the usual common sense and naturalistic concepts. Because it presupposes meaning and intention, "world" is not identical with the totality of entities, as common sense would have it, nor with the cosmos studied by natural science. Common sense and science treat what Heidegger calls "world" as a system of subjective attributions with no ontological significance. But Heidegger regards world in his sense as ontologically fundamental and claims that our ordinary common sense and natural science are founded on it.

Heidegger develops his concept of world as an "existentiale," that is, as a universal feature of being in its relation to human being. World is a "category" in the

Aristotelian sense, but a category dependent on the human subject. The universality of such categories overleaps any particular cultural limitation to define the human as such in its relation to being. What is generally called culture enters this picture under the name "*das Man*," the "they."

Heidegger's analysis of worldhood is intended to overcome the subject/object ontology he identifies with the tradition of modern philosophy. The world is referred ultimately to a subject that Heidegger calls "*Dasein*" to avoid confusion with consciousness, the idealist subject. Under the influence of Dilthey, Heidegger originally called this subject "factical life." This designation indicated the two features that distinguished his concept of subjectivity from the traditional one. On the one hand the subject is not to be conceived as a spiritual entity, a substantial thought, a *cogito*, but as a living being, hence a being essentially connected to its surroundings. On the other hand life must be grasped from the inside as a way of being rather than from the outside as an object. "*Dasein*" continues to signify this lived relationship of life to itself.

Being and Time explains the concept of world on the model of the workshop and its tools.[4] The tools are linked together by their relations in the work. They do not stand alone. Their functionality is granted by their place in the whole of the workshop. In sum, artifacts serve in a context. He writes,

> Now in the production of equipment the plan is determined in advance by the serviceability [*Dienlichkeit*] of the equipment. This serviceability is regulated by anticipating what purpose the piece of equipment or indeed the machine are to serve. All equipment is what it is and the way it is only within a particular context. This context is determined by a totality of involvements [*Bewandtnisganzheit*] in each case. (Heidegger 1995a, p. 215)

This totality is a system of references between the entities in *Dasein*'*s* world. Heidegger calls this system "significance" (*Bedeutsamkeit*) and treats it as an open space of meaning within which particular usages are enabled.

The workshop example illustrates this unitary subject-object, which he calls "being-in-the-world." *Dasein* and its tools belong together. "Being-in-the-world" consists in the connections between technical artifacts and the ordering role of the human being at the center of the technical network.

Heidegger also defines world as "beings in their accessibility" (Heidegger 1995a, p. 199). By "accessible" he means understandable *as*, taken *as*, enacted *as*. Thus the chair on which I sit is not simply there as an object but is treated by me *as* a chair, that is, as intended for sitting. No such relation to the chair is possible for the papers I stack on it temporarily in my preparations for leaving the office. Those papers are supported by the chair, but not *as* a chair. *Dasein* establishes a different type of relation from the causal relations among things, a relation of meaning. In this sense, then, worlds are existential situations, not collections of things. Perhaps the closest our everyday talk comes to Heidegger's own usage is in expressions such as "the world of the theatre," "the Medieval world." Such worlds are not merely subjective

---

[4] The principal discussion of tools is in Heidegger, *Being and Time*, part 1, section III; the distinction in world relations is developed at length in Heidegger, *The Fundamental Concepts of Metaphysics* : *World*, *Finitude*, *Solitude*, part 2, chapter 3.

but nor are they the sum of the things they encompass. They are essentially related to *Dasein* without being reducible to it.

Dasein's principal characteristic is concern with its own being. This concern is played out in the constitution of an environment distinct from the totality of nature as understood by natural science. Nature as an object of knowledge includes much that is of no concern to the living subject. Those irrelevant aspects are discovered in objective contemplation but are not part of the original world constituting relationship. That relationship consists in the network of functional references that enables *Dasein* to get around and to further its aims. "*Das Man*," culture, sets the terms of the references.

The common sense view of the difference between the cognitive stances of the actor and the objective observer comes down to a difference in focus. In the first case the focus is on what ties the object into the network of references. Heidegger's workshop is full of objects understood exclusively through their functional properties. The hammer is hard, has an appropriate weight in the hand, and can be swung in a specific arc at the nails to which it "refers" in performing the work for the actor who wields it. It is, says Heidegger, "ready-to-hand." It is not composed of iron atoms nor is it made in a certain factory on a certain date, nor was it formerly owned by Mr. X or Ms. Y. Those objective "present-at-hand" attributes are of course accessible to the subject in principle but they are not focused in the active employment of the artifact; they are not part of the subject's world.

The "sight" associated with action is not explicit propositional knowledge but is what we now call "tacit" knowledge, practical know-how, "circumspection" in the English translation (Heidegger 1962, p. 98). Correspondingly, the subject of this knowledge is to be understood through its involvement in the technical network. It is not a separate *cogito*, a pure mind, but is an active being enmeshed in a world of objects with which it is essentially concerned and to which it is essentially connected.

Although he is not interested in developing a theory of function as such, his argument illuminates important aspects of a phenomenology of function and invites completion along lines compatible with his contribution. His essential insight is the concept of "involvement." He says that entities must be "freed" for their involvement through entry into the system of references. The entry of an entity takes place through establishing connections to those attributes that make it available for the referential relation. Today we might call this the "affordances" of the object. Heidegger develops this concept in an unusual account of production that has suggestive implications for the understanding of functionality.

In *Being and Time* Heidegger is more interested in explaining everyday action than technical production. His comments on production are accordingly quite brief, but they do clearly distinguish the materials of production from present-at-hand nature. He points out that production incorporates natural materials into technical objects, but he rejects the notion that these materials are identical to the objects of natural science. They belong to the world even before they are worked up into a specific technical object for a specific purpose (Heidegger 1962, pp. 100–101). Exactly how they belong Heidegger does not say.

The closest he comes to a theory of production in the early work is an analysis of Aristotle's concept of *dynamis* in the *Metaphysics* (Heidegger 1995b). On phenomenological terms, the materials of production are "freed" in some non-specific way that invites a variety of uses. The selection of some among those possibilities would, in removing the ambiguity of the materials, remove them from the context in which they are originally revealed in their indeterminate multiplicity and reduce them to their useful qualities in a specific context of use.

Thus the materials are not objective things in the full sense, nor are they already technical objects; they belong to the world through their *potentialities*, i.e. through what they can "bear" or "tolerate" (*pathein*), the many referential relations in which they *could* be involved even before they enter a specific production process. The production process that realizes one among those potentialities is a narrowing down, a limitation (*peras*), through incorporation of the material into a specific network of references. Heidegger concludes that production actualizes the *telos* of its materials. Employing the example of pottery, he writes, "With the transformation of the clay into the bowl, the lump also loses its form, but fundamentally it loses its formlessness; it gives up a lack, and hence the tolerating here is at once a positive contribution to the development of something higher" (Heidegger 1995b, p. 74).

A tree can serve as an example of the implications of the theory. Even while it grows it belongs to the world as a potential source of useful objects, such as a telephone pole, lumber, paper, etc. The reduction of the tree to a single potential begins by removing it from the growth setting in which it was potentially referenced in these various ways, and stripping its branches and bark. This is done in terms of a choice of a specific referential system, for example, one that involves the tree as lumber for building a house. Certain useful qualities of the tree are privileged over others. Those qualities tie the lumber into the referential system of carpentry, its tools, procedures, and designs. Further references are supplied by the detailed specifications of the particular house to be built. Ultimately a product is realized through imposing successive limits on the potentials of the growing tree and in so doing actualizing a house.

There is an ambiguity in Heidegger's theory of function. This is clear in the example of the house. From his descriptions of tool use one might think that only hammers, nails and lumber are involved, but we know that the referential framework of a house includes much more than this bare technical minimum. In the final design the lumber acquires qualities it would not otherwise possess, such as aesthetic features, conformity to rules of the trade, and so on.

The boards in the American construction system are posed horizontally, whereas in Scandinavia they stand vertically. The rules of the trade differ as does the aesthetic effect. There are also legal regulations to which the house must conform, the building code determined by local legislation. These additional references are normative mediations of the construction process which intervene at various stages to further narrow the range of possibilities. They compensate for the simplifications that enable the materials to appear as materials for this or that specific project. All this would be included in what Aristotle calls "form" and what we might call "cultural meaning." Through these mediations the final product takes its proper place in a social context, a cultural system. Functionality in the narrow sense of the term to which we are accustomed is an abstraction from this always present, richer system of references.

In everyday non-phenomenological language this amounts to removing the object from its natural context, reducing it to its useful properties, systematizing it in a new humanly created context, and mediating it in terms of norms that correspond to qualities it did not possess in nature. But Heidegger resists this explanation because it presupposes an understanding of the object as a thing prior to its involvements in a world. For him practical relatedness comes first and is fundamental. The decontextualization and reduction operates within the world, not in a relation of the subject to objective nature. Similarly, the process of systematization and mediation is not an engagement with objective things but addresses objects only insofar as they already belong to the world. The difference between these two accounts is of great significance to Heidegger but less so for a theory of function. The four basic operations essential to the functionalization process, *decontextualization*, *reduction*, *systematization* and *mediation*, are performed on the terms of both accounts, naturalistic and phenomenological.

This concept of functionalization can be articulated with the notion of assigning a function in the use-plan conception of Houkes and Vermaas. In their framework a functional assignment presupposes the belief that the object possesses the causal properties necessary to perform the function. What are those properties? Clearly they are not the sum total of what an objective view of the object would reveal, nor are they the product of disinterested observation. In making a functional assignment, the subject must know about those properties of the object that are relevant to its technical operation. That small subset corresponds on the side of "belief" to Heidegger's concept of "circumspection."

For example, the individual who assigns the function of hammering nails to the hammer must believe that it is hard enough to do the job. But that belief is contingent on understanding the hammer exclusively in its belonging to the workshop, as a carpentry tool, as opposed to understanding it in relation to the infinite variety of contexts in which it participates as a thing. The belief that enables the assignment focuses on the hardness of the hammer as the condition of its functionality to the exclusion of an infinity of other properties. The positive quality of the hammer as a technical object is thus also a limit. In actual use the same limit makes the hammer available for the user's activity. Whether one calls that limit the constitution of a world (Heidegger) or a belief about things (ICE), it is essential to the nature of function.

16.3 Technification in Heidegger

The analytic concept of "belief" in ICE is vague. It is stretched to cover the teleological understanding of tools and the objective knowledge underlying modern technology. Heidegger's early work relates objectivity to science, but not to technology. It is only after World War II that he develops a full fledged theory of technology. That theory is an account of how science depends on and supports a practical intent to control and dominate nature. Heidegger interprets this technical relationship to reality as an ontological clue, just as he did in his earlier analysis of tools. But technology reveals a very different reality.

Modern science, he claims, is essentially technological. It sets out a "ground plan" of being as a lawful order of facts. This is the constitution of a realm of objects subject to scientific explanation. On that basis it makes predictions that guide the technological transformation of what is. Technology is thus the opposite of world in *Being and Tim*. The world is a totality of ready-to-hand things engaged with *Dasein*, By contrast technology is a representation of present-at-hand things before a cognitive subject. Technology is the triumph of detached representation of things, and of the subject of such representation, over the involved stance of the acting subject described in the early work (Heidegger 1977).

Technology does not construct a world in the sense in which Heidegger originally understood that concept, but de-worlds its objects and reduces them to raw materials in a process planned in advance in view of predictable results. Modern technology "enframes" man and nature. It "challenges" nature and makes "unreasonable demands" on it. Things no longer realize potentialities within a world but are stripped bare of qualities, of their very thinghood, to take their place in a technological system. They are no longer objects in the sense of having a being that confronts us (*Gegenstand*); they have become mere resources (*Bestand*).

This Heideggerian theory of technology treats functionality in the modern context as the loss of substantial reality. Functionalization leaves only matter and energy (Zimmerman 1990, p. 212). Things are reduced to what Heidegger calls "standing reserve." They are extracted from their surroundings, stored up, moved, and transformed to perform unnatural feats. What this amounts to on the terms of the earlier analysis is the loss of the complex systematizations and mediations that situate objects in a world, the meanings and norms imposed along with the manipulations in which technical practice consists. What remains is the bare minimum required to exercise control. "The outstanding feature of modern technology lies in the fact that it is not at all any longer merely 'means' and no longer merely stands in 'service' for others, but instead…unfolds a specific character of domination" (quoted in Zimmerman 1990, p. 214).

Heidegger's theory of technification provides still more specification of the beliefs associated with functional attributions. As noted above these beliefs concern only those properties of the object that are relevant to its operation in its technical setting. Heidegger's late work adds to this limitation the specific property of being law-governed. The relevant beliefs must include the idea of a law under which the object can be made to serve technically. This explains the privileged role of causality in the beliefs associated with functional attributions in modern societies. By contrast the role of cultural meaning and the significance of the lifeworld described in Heidegger's earlier work are eclipsed in modern times by an implicit ontology that takes the nature of natural science as the only real. The theory of technification also offers a hint of a theory of modern technical subjectivity, emphasizing the detached cognitive standpoint from which a plan is constructed.

Heidegger's negative evaluation of modern technology presupposes an implicit critical standard, the teleological view of nature underlying his early theory. But he does not defend the earlier view explicitly in his later work. He never advocates the teleological concept of production even as he criticizes modern technology. To do so would be to regress to a premodern conception of *poiesis*, and Heidegger does not believe it possible to go backward in what he calls the "history of being" (Heidegger 1977, p. 136). But the way forward is obscure.

There is a further difficulty with Heidegger's later theory. It is unclear whether he believes that the simple attribution of function to an object changes its essence, or if that attribution initiates a material process of transformation. He argues for example that the hydro-electric plant placed on the Rhine river transforms the river into a resource (Heidegger 1977, p. 16). But is it the simple functional ascription of the river that has this effect or the actual material intervention represented by the power plant?

Contemporary critics of technology inspired by Heidegger generally maintain the ambiguity, but offer more concrete accounts of the material transformations that objects and human relations undergo when they enter the functional realm. Technification is a process with specific characteristics that flow from the nature of functionality. The cognitive narrowing and limitation associated with a modern functional perspective cuts off dimensions of objects and persons that are worthy of preservation and respect, but modern culture privileges the causal characteristics of artifacts above all else. Albert Borgmann gives the example of the family dinner, a ritual occasion shattered by the reduction of dining to a functional minimum through the mere ingestion of micro-waved or fast food (Borgmann 1984, p. 105).

Such arguments imply that the spread of a functional standpoint beyond certain purely technical bounds is a spiritual catastrophe. The theorists plead for limitation of the functional realm at least as it is realized technologically (Böhme 2012, p. 194). This plea responds to the radical simplifications involved in constructing the technological object. Those simplifications are incompatible with many other relations to objects that sustain them in their thinghood and worldly character. The problem from this standpoint is thus not the existence of function but its imperialism in modern societies.

This type of critique depends on a teleological interpretation of the *human* context from which a technical function is extracted. Thus the focus shifts from technology itself to the re-ordering of human relations and the associated objects around technical mediations. In this way the critique of the generalization of functionality in modern societies is saved from the reactionary nostalgia that sometimes threatens Heidegger's own discourse. But a social critique is substituted for the ontological intent of Heidegger's theory. We are squarely in the domain that the early Marxist Lukács explored with his theory of reification.

16.4 Lukács's Philosophy of Technology

Georg Lukács was a Hungarian philosopher and literary critic who wrote most of his work in German and participated in his early years in the German cultural world that also shaped Heidegger's philosophy. However, the politics of these two philosophers could not be more different. Lukács became a Marxist at the end of World War I and in 1923 published a classic work of Marxist philosophy entitled *History and Class Consciousness*. In this book Lukács put Marxism in touch with contemporary sociology and Hegel. The result is a remarkably original reconstruction of Marxism as a critique of modern rationalized society. Lukács had a profound influence on the Frankfurt School and on what Merleau-Ponty called "Western Marxism" (1955).

It is interesting to note that Lukács's *History and Class Consciousness* anticipates Heidegger's later theory of technology, first proposed in a speech in 1949. Both argue that modernity (in Lukács's case *capitalist* modernity) is characterized by the tendency to functionalize the entire world in terms of a scientific-technical conception of the law-governed order of nature and society. Like Heidegger, Lukács contrasts the concrete objects of premodern societies with the stripped down products of modern technology (Lukács 1971, pp. 97 and 236). No doubt Lukács's economic focus is due as much to the less prominent role of technology before World War II as to his Marxism, which offers hope that a socialist society can radically alter the role of technology. Heidegger treats all modern societies as similar, after the demonstration of the absolute power of technology in the War and the betrayal of the promise of the Soviet Union.

Lukács was no more interested than Heidegger in contributing a philosophy of function but his reflections are rich in implications for such a philosophy. As noted at the outset, functionality is a two-sided affair, referring to both the subject and the object. Heidegger's theories of worldhood and technification have been helpful in thinking about the objective phase of functionality, but we have not yet explored its subjective phase. Lukács's theory of reification is useful for this aspect of the theory of function. Furthermore, Lukács's theory makes explicit the technical character of the whole technosystem, including administrations and markets.

Lukács notes the similarity between the economic laws of political economy and the laws of nature discovered by natural science. He argues that the capitalist economy is actually law governed as though part of the natural world. It is a kind of second nature, resembling the first nature insofar as it too is subject to technical manipulation on the basis of its laws. Lukács thus treats the economy as a realm of technical action. He writes,

> What is important is to recognize clearly that all human relations (viewed as the objects of social activity) assume increasingly the form of objectivity of the abstract elements of the conceptual systems of natural science and of the abstract substrata of the laws of nature. And also, the subject of this 'action' likewise assumes increasingly the attitude of the pure observer of these—artificially abstract—processes, the attitude of the experimenter. (Lukács 1971, p. 131)

Even though the economic and social system comes to resemble the nature of natural science, there is a difference of principle between them. In the case of nature, the laws are matters of fact, whereas the laws regulating the capitalist economy are the product of human actions, specifically, the technical manipulations through which individuals pursue their economic interests. Lukács calls the capitalist economy “reified” in the sense that it appears as a thing when in reality it is an unconscious process of human relations. However, the thing-like appearance of the economy is not an illusion. It has real consequences to the extent that the appearance motivates people to perform the type of actions that reproduces it.

The circular relation between economic laws and the technical manipulations which both generate the laws and take advantage of them is fundamentally different from the case of nature in which the technical manipulations have no role at all in generating the laws. The individuals can break out of the circle through cooperative action to change the system. This de-reifying practice is synonymous with the proletarian revolution. It is not a technical manipulation of the economy in accordance with its laws but the overthrow of those laws through a transformation of their practical basis.

From the standpoint of a theory of function, Lukács’s contribution is a conception of the functional subject. This is an individual subject that stands at a distance from its objects. It is autonomous, uninvolved in the objects it functionalizes. Lukács calls its practice “contemplative” in the sense that it does not aim to change the nature of its objects but only to manipulate them. Manipulation posits the law of the object as fixed and unchangeable in order to control superficial features of things that stand under the law. This manipulative orientation blocks feedback from the objects. The subject posits itself as outside the system in acting on it.

In economic terms, this amounts to taking up a position with respect to what the objects will become in any case under the control of their laws. This is most obvious in the case of the stock market. The action of the “contemplative” subject consists in buying stocks it expects will increase in value. The subject positions itself with respect to the lawful development of the economy rather than attempting to shape that development. For Lukács this is the “model” of practice throughout capitalist society (Lukács 1971, pp. 83 and 98). In a technologically advanced production process the worker stands in a similar contemplative relation to the self-acting machinery he or she operates. The bureaucrat too operates manipulatively under the rule rather than acting to change the rule.

This is a narrowed relationship to which corresponds a narrow subject. The functional subject is stripped bare of personal qualities that would only interfere with successful manipulation. What is lost in this narrowing is far richer than Heidegger imagines, not simply self-knowledge as the site of revealing, but concrete human qualities and needs. Despite this critical perspective, Lukács is not opposed to technical practice in principle; it will after all be required by any modern society, including a socialist society. But he believes that a world and a subjectivity narrowed down to the measure of technique cannot fulfill the human potentials (Lukács 1971, p. 6). Once again, as in Heidegger, the universalization of the technical is the problem, not technology as such.

Lukács develops this argument in terms of the tension between reified form and living human content. Technology and other reified systems impose a formal structure on life and behavior. The form, reification, consists in fragmenting and isolating social objects as though they were self-subsistent things, like things of nature, related only externally, causally. Under capitalism that structure is oriented toward the production of profit. Human beings, in the fullness of their existence and needs, are forced into the structure without regard for consequences. This process both generates a potential and represses it. What human beings can become is laid out in their relations within the reified system, but only as potentiality, not as actuality. This tension between form and content motivates the revolution. Thus once again, as in Heidegger, the concept of potentiality provides an alternative to technical domination. This alternative can be realized where functional relationships are set in place within the higher level of a collective political subject. But whereas for Heidegger potentiality lies in the Greek past, for Lukács it awaits in the communist future.

The analytic theory of function remains at the level of individual technical action on nature and so does not venture onto the terrain of the social arrangements which set the conditions for individual action. The theory conforms more or less to Lukács's concept of contemplative action. The functional subject's beliefs concern laws over which it has no control and which it can only use, not change. This seems self-evident in the case of nature, which provides most of the examples in the analytic theory. These examples are appropriate for a subject conceived as an individual engaged in a single round of action based on a conscious goal.

But in the real world functionalization extends far beyond the kitchen utensils, guitars, and cars that provide examples for the theory. As Heidegger and Lukács argue, technological and bureaucratic systems structure much human action and cannot be regarded as mere means. They shape and damage human life even as they serve. And Lukács is not wrong to view economic action in terms of functional relations. Entering a store, the buyer confronts the salesperson in his or her function as an economic agent. Of course sympathy may arise between seller and buyer, or antipathy for that matter, exceeding the limits of a functional relation. But in the normal case the two parties to the transaction "use" each other for their own ends in accordance with an intention and associated beliefs. The point is not that this is inherently bad, but that multiplied millions of times over it constructs a coherent system, the capitalist economy, which compels the adoption of a technical stance in more and more of social life, thereby creating a world characterized by inequality and domination.

As in the case of Heidegger's concept of world, so with Lukács's concept of the functional subject and its practice, the focus is excessively narrow. I note two such limitations. First, Lukács omits the general consequences of technical practice for identity to focus exclusively on the role of reification. But technical practice shapes subjects in far more complex ways than he allows. And second, Lukács has no concept of social imagination with which to understand original initiatives, creative action. But we know that creativity drives capitalism forward and would be essential to any fundamental social change.

The contemplative stance of the reified subject has a paradoxical effect. At the same time as the subject withdraws from involvement, the technical relation to the world determines its identity. Thus in avoiding causal feedback from its objects, the subject of technical practice simply shifts the interaction to the higher level of meaning. Lukács mentions only two cases but others can easily be imagined. He argues that the journalists and bureaucrats, individuals who invest their personality in their work, form an identification with the reified system. These middle class individuals have a sense of self and beliefs about the world in which the limitations of capitalism have become personal limitations of character and understanding. By contrast, workers' identity cannot be formed by their work since the production process is so much more profoundly reified. They retain a certain independence since only mechanical physical gestures are required of them (Lukács 1971, pp. 99–100). Insofar as their participation in technical practice is identity forming, this is through their realization that they are more than the social role to which they are condemned. They are thus capable of initiatives that challenge the system as a whole. For Lukács this is the origin of class consciousness.

Now, for class consciousness to lead to positive initiatives, imagination is required but Lukács does not discuss this aspect of the revolution. He underestimates the role of the imagination through which individuals may transcend the narrow limits of their position in the economy and take unprogrammed initiatives in which new functions are discovered. This is as true of the middle class individuals he criticizes for their reified consciousness as of the workers whose ability to transcend their situation he attributes to the poverty and injustice they suffer. Initiative and imagination are powerful forces under capitalism, though often exploited or repressed, and their importance must not diminish in a socialist society.

Such initiative is not to be conceived as *ex nihilo* creation but rather as rooted in conditions which can be dereified in specific ways that release the potentialities those very conditions create and block. This notion of dereification bears some resemblance to Heidegger's description of authentic resoluteness. This is release from the pre-given references, the cultural system provided by *das Man*, to an original decision (Heidegger 1962, p. 345). Indeed, initiative in a technologically advanced society must have an innovative character that breaks with sterile conformism. But neither Heidegger nor Lukács applied this insight specifically to technological change.

In his early work under Heidegger's supervision, Herbert Marcuse developed the implications of political initiative for revolutionary communism.[5] His unusual synthesis of Heidegger and Lukács draws together the concepts of authenticity and revolutionary praxis. In these early writings Marcuse is on the verge of a positive concept of initiative. His late work completes the picture, relating revolution to the imagination of alternative social and technological institutions. Technology now enters directly into the theory as an object of imaginative reconstruction. Thus Marcuse can be interpreted as theorizing the aspects of initiative and imagination

---

[5] His late work relates it to the imagination of alternative social and technological institutions (Marcuse 1964, pp. 228 and 239–240; Feenberg 2005, chapter 5).

that are underdeveloped in Heidegger and Marcuse, but only within the specific historical framework of Marxism, and not in a general theory of functionalization (Marcuse 1964, pp. 228 and 239–240, 2005, pp. 31–32; Feenberg 2005, chapter 5).

This brief discussion of Lukács's theory of reification in relation to Heidegger and Marcuse highlights four aspects of functional subjectivity. It shows the autonomy of the subject role with respect to the objects of technical practice, while also hinting at the consequences of that practice for identity. It contrasts positioning as contemplative practice with initiatives transcending the reified framework.

16.5 The Double Aspects of Technology

I will turn now to my attempt to synthesize the arguments of Heidegger and Lukács in a framework for understanding technology. I call this framework the "instrumentalization theory." It is the result of a long evolution that to a certain extent parallels the "dual nature" project exemplified by the analytic use-plan theory of function (Kroes and Meijers 2002). My own project began in 1975 with an invitation to a conference at the Villa Serbelloni on "Technology and Communist Culture." This was a first opportunity to think seriously about the nature of technology. As a democratic socialist I was faced with a dilemma. The then widely held deterministic theory of technology argued that democratic control of the economy was incompatible with technological "imperatives." I rejected this conclusion yet also rejected the instrumentalist notion that technologies are value-neutral means. That technology has social impacts I had no doubt, but I argued that a democratic socialist regime could develop technology with different social impacts compatible with its values.

A footnote to the conference proceedings summed up the basis of this argument I have been developing ever since.

> At the Bellagio Conference I suggested a terminology with which to distinguish the neutral from the socially determined aspects of what might be called the 'technosphere.' I would reserve the term 'technique' for specific technical elements such as the lever, the wheel, the electric circuit, and so on, all of which are in themselves neutral applications of objective knowledge of nature. These elements are like the vocabulary of a language; they can be strung together to form a variety of 'sentences' with different meanings and purposes. 'Technologies,' defined as developed ensembles of technical elements, are greater than the sums of their parts. They meet social criteria of purpose in the very selection and arrangement of the intrinsically neutral units from which they are built up. These social criteria can be understood as 'embodied' in the technology and not simply as an extrinsic use to which a neutral tool might be put. (Feenberg 1977, p. 114)

In sum, technology has a foot in two worlds, a world of rational structures and a world of human intentions that organize those structures for a purpose. I later called this a "double aspect" theory of technology with the implied reference to double aspect theories of the mind/body relation (Feenberg 1991, pp. 78 and 83, 1992, p. 311). When Descartes separated mind from body, he relegated the body to the mechanical realm. I wanted to block Cartesian dualism in the understanding of the

Table 16.1 Instrumentalization theory

	Causal functionalization	Cultural functionalization
Objectification	Decontextualization	Systematization
	Reduction	Mediation
Subjectivation	Autonomization	Identity
	Positioning	Initiative
Cognitive relation	Nature	Lifeworld

Adapted from Feenberg (1999, p. 208)

mechanical itself. A purely mechanical explanation of technology threatened a return to naïve instrumentalism or technological determinism.

In the remainder of this section I will outline the contribution of critical theory of technology to an understanding of functionality. The treatment differs from analytic approaches to the same theme; rather than conceptual analysis, I propose a focus on social aspects of the phenomena. The analytic discussions emphasize cognitive aspects: the ascription of a function rests on a belief that the materials are appropriate to fulfilling their function. Drawing a more complete picture of the functionalization is a descriptive phenomenological task. The theory is action-theoretic not just in attending to beliefs and intentions of actors, but in the sense that it describes the nature of the subjects and objects of technical action. Thus instead of analyzing the *concept* of function in detail, the argument concerns *the subjective and objective conditions of functionalization as a social process*.[6] Understanding that process has normative implications insofar as it opens up possibilities of change foreclosed in the deterministic theories of technology that still have great influence in policy circles.

The above table sums up these conditions. In the remainder of this section I will explain the terms of the table (Table 16.1).

The causal level concerns the construction of objects and subjects as nature, again in a "meta" sense, that is, as subject to rules or laws that regulate their behavior as materials. The causal level corresponds to a mental capacity possessed by a few higher animals and especially by human beings. It is a particular way of relating to the world that makes technical usage possible. The cultural level concerns the meanings artifacts acquire in the lifeworld to which they belong. This is a uniquely human aspect of the technical for only human beings are capable of both technical mentality and cultural creation.[7] These meanings are not merely ascribed after the

---

[6] Ted Cavanagh (2008) worked out a useful example of the application of the theory to building construction. The terms of the theory at that time were somewhat different. I called "causality" and "culture" "primary" and "secondary" instrumentalization in earlier versions. This led to confusion between secondary instrumentalization and processes of reinvention or creative appropriation which occur after the technical artifact is released on the public whereas my intent was to describe complementary aspects of all designing.

[7] Gilbert Simondon has written an important article on the technical mentality, but he mistakes the cultural level of the technical artifact for a distraction from an imagined pure technique. See Simondon (2009).

causal level is set in place, but guide the choice and configuration of the causal level. The specific cultural assumptions may be universally shared as in primitive societies, or they may be imposed by social forces. In the latter case these social forces are exercised directly by influential groups, "actors" in the terminology of social constructivism. In every case a combination of causality and culture is involved in creating the framework within which function is perceived by the makers and users of technical artifacts.

The causal and cultural layers of the design process are analytically distinguishable phases. They are visible from different perspectives but cannot be separated and laid out side by side. One phase involves the conditions of the relevant rule-based or causal foundation of the functional ascription, and the other posits the guiding cultural meanings that determine relevance and signify the object. The two phases together identify potentials that are selected and combined in the realization of the design. The layers interpenetrate in the sense that a causal relation is realized only insofar as it has been invested with cultural meaning. In this respect the instrumentalization theory goes beyond the use-plan theory of function.

Except in the case of the very simplest of artifacts, the functional investment of an object involves more than a subjective intention; it determines a choice of components and the relations between them, that is, a design. Realization in a design can take many different paths. There is no universal rule under which to make the choice of functions from among the infinite possibilities, although all such choices must conform to causal principles. It is this contingency of design which opens the way to a politics of technology as I will argue in a later section of this chapter.

This initial distinction in layers can be analyzed further into the objective and subjective conditions of design. In the phenomenological language of Husserl and Heidegger we would say that the "object is revealed as…" and the "subject constitutes itself as…." We have seen that for Heidegger the object is "freed for" entry into a world. In ordinary language this means roughly that the object is envisaged under the aspect of its technical potential and the subject adopts a technical attitude toward it, that is, again in Heidegger's terms, acts toward it out of its concern with its own identity or "being" as he calls it.

The ascription of function requires more than a general belief in causal appropriateness; it also requires a specific type of cognitive operation, a technical mentality that goes beyond the immediate form of the object and reveals it in the light of its technical potential in a specific cultural context. In the instrumentalization theory the initial correlates of this operation on the side of the object are called decontextualization and reduction. Technical potential is uncovered through isolating the object from its natural context and reducing it to its usable qualities. On Lukács's terms, the object is reified.

The object must be processed in order to be incorporated into an artifact. The processing does violence to the object in its original state, transferring it from nature to lifeworld. As we have seen Heidegger conceptualizes this process in two different ways corresponding to different stages in the development of technology, either as the actualization of a potential or reduction to raw materials.

Realization takes place in a system of technical and social contexts that guide decontextualization and reduction. This involves systematization and mediation. The technical object is taken up in the system of references that Heidegger describes as a world. This system involves more than simple causal relations. The object cannot enter the social world without acquiring a social meaning. This concept of meaning refers to the many associations of the object with other aspects of the culture, including aesthetic and ethical mediations of the design. The object may be signified socially by its price, as Lukács argues following Marx, but it has a use value as well. At that level the object belongs to a lifeworld in which it is imbricated with many other aspects of nature and human life. Technical objects thus not only lose qualities as they are reduced, but acquire qualities as they are integrated to a social world.

The instrumentalization theory identifies a basic technical attitude that allows objects in the world to be envisaged as artifacts or components. This attitude constitutes the technical subject. It has two aspects I call autonomization and positioning. The basic attitude is autonomous in the sense that it precludes sympathy or identification, the attitudes that are associated with human relations, i.e. relations with another subject. In the technical relation the subject is not involved in reciprocal interactions. It is protected from feedback from its objects. The point of technical action is to change the world, not the technical subject.

As Lukács argues, the subject does not strive to create something entirely new but takes up a position with respect to what the object is and will become, a position that opens up its useful potentials. This is a manipulative attitude, one that seeks control of the object through an understanding of its properties, the "law" of its movement.

Correlated with these causally related functionalizations are two other aspects on the cultural side of technical activity, identity and initiative. The technical subject acquires an identity through its association with its objects. This may minimally signify that the subject can be described by its use of its objects as when we say of persons driving that they are drivers. Where there is extensive and long term technical work, professional identities are established by repeated functional involvements. Both Heidegger and Lukács hint at this concept of identity in showing the intimate connection between technical subjects and objects.

In every case but most importantly in professional activity, the technical subject exercises a certain freedom or initiative in the discovery of the potentials of its materials. The scope for initiative varies but it is an inevitable result of placing individuals in a technical relation. At a minimum, the initiative is defined by the range of activities enabled by the design of the object. But it may go beyond the normal range and explore potentials unsuspected by the designers. This is the basis of the creative appropriation or reinvention of technologies by users. This concept of initiative appears too in Heidegger and Lukács, although it is limited by being tied to notions of authenticity and revolution.

16.6 Political Implications

In conclusion I will return to the question of the political implications of the philosophy of technology as they now appear to me many years after my initial work on the question of technology. My original socialist argument with determinism has not been abandoned, but it has been supplemented by a broader political theory of technology capable of interpreting social movements such as the environmental movement.

The instrumentalization theory is intended to open up the imagination of the future to a possible transformation of industrial society. It does so by showing how the level of technology that is causally determined is contingent on the evolution of culture. The deterministic argument depends on conflating the binding rationality of the first level with the contingent social choices involved in the second. It claims that the causally coherent pattern of technology is the product of a process of development determined by scientific knowledge, as though no cultural mediations intervened between scientific discovery and the making of artifacts. The trajectory of development can then be projected independent of changes in social organization and culture. A technocratic politics corresponds to this approach.

The instrumentalization theory is central to my approach because it blocks the two reductionisms that distort the understanding of technology and lead directly to this technocratic politics. On the one hand, there is a tendency to reduce the cultural level of technology to the causal level. This takes the form of explaining what are in fact culturally imposed guidelines in the specification of technology as consequences of those very technical specifications. For example, if the Internet is used for human communication that is *because* it is technically specified to support that usage, not because a process of social and cultural change guided the development of the specification. This makes it seem as though technology is purely technical, its form and workings entirely independent of society. On the other hand, there is a related tendency to reduce the social world as a whole to its technical underpinnings, as though technology, as a material "base," could determine the "superstructural" features of society. This tendency, derived from a certain interpretation of Marxism, has technocratic political consequences in both standard strategies of modernization and in the failed communist experiment.

The instrumentalization theory provides an alternative to reductionism by showing the true relations between the different aspects of the technical domain. It shows the role of culture in the configuration of the causal aspect of technology from which it can only be distinguished analytically. This counter-argument to determinism opens up a role for social struggle and politics in technological development.

The instrumentalization theory depends on the constructivist notion of "under-determination." There are always alternative designs with different implications for social groups, hence designs are "under-determined" by purely technical considerations. As Pinch and Bijker write, "The different interpretations by social groups of the content of artifacts lead by means of different chains of problems and solutions to different further developments…" (Pinch and Bijker 1987, p. 42). The point is that the "content," the actual causal concatenation that goes into the design, is socially contingent. The instrumentalization theory attempts to explain the fundamental relationship between technical and social aspects of technology in an underdetermined technical world.

This conception has liberating political implications. The future is no longer bound in its basic design by a continuation of the trajectory of the existing technologies. Indeed, even the existing technology is revealed as contingent and subject to radical transformation. Technology can be placed on a different trajectory in conformity with different values. Examples of incipient changes of this sort are not hard to find.

Consider deskilling, the reduction of the skills required to operate the ever more powerful machinery of industry. This was my principal concern at the 1975 conference that started me on the way to a philosophy of technology. At the end of the deskilling trajectory human beings are simply replaced by intelligent machines. This trajectory has two effects: it increases the availability of consumer goods even as it erases the vocational identity of the workers who make them. Can we imagine a different trajectory of development in which the advance of technology opens opportunities to apply intelligence at work instead of replacing it with mechanical substitutes? There are in fact experiments in alternative types of development, but how are these to be evaluated? Can they be dismissed out of hand for violating the supposed imperatives of technology? The instrumentalization theory makes no predictions, but it offers a stimulus to the imagination of a socialist alternative where the deterministic theory shut down the very idea of a different future with its mantra of inevitability.

Environmental problems raise similar issues. We are told that environmental reform is incompatible with industrial progress, that we must burn fossil fuels and pollute the planet to enjoy a modern way of life. But again the deterministic premise is challenged by the many initiatives taken in recent years in response to environmental problems. Innovation and regulation stimulated by social protest are transforming industry. How we evaluate that process depends on our philosophy of technology. An alternative technological system is possible on the terms of the instrumentalization theory in contrast with determinism.

The evolution of the Internet offers the best evidence for the instrumentalization theory. The system was originally designed for a completely different purpose than the ones it now serves. Its technological trajectory was set by the military and the research community and focused on the efficient use of computing resources. But once the Internet was released on a different public, the usages changed. It turns out that what these users want from the Internet is opportunities for human communication. A tremendous flowering of new social forms and usages has changed our understanding of both the computer and society. Deterministic approaches take these changes so much for granted they appear as the inevitable result of the development of computing, but the instrumentalization theory allows us to focus on the concrete initiatives that have led to the present situation.

In sum the instrumentalization theory is politically significant not because it advocates or supports any particular politics but because it makes politics thinkable in a world completely penetrated by technologies and technical systems. In the absence of a theory supporting a social understanding of technology technocracy beckons as the only rational organization of a society based on scientific knowledge. Philosophy of technology is called to challenge this conclusion with an account of the reciprocal relation of society and technology.

16.7 Conclusion

In this chapter I have contrasted the instrumentalization theory with the analytic theory of function, specifically the ICE theory introduced by Houkes and Vermaas. ICE explains the relation of the individual to technical artifacts and natural objects of technique. This is an important context of human action, however, it is always contextualized by a social world that the instrumentalization theory brings into the picture. Thus the approaches are complementary. The use-plan theory can be conceived as an abstraction from a larger social context explored in the instrumentalization theory.

Together the two theories offer a powerful approach to understanding the concept and nature of function, however, neither theory is complete. The instrumentalization theory as formulated here lacks an account of the structure of the interactions between the two levels of functionalization it describes. Are these interactions purely contingent or do they exhibit general features that ought to be theorized to complete the theory? The ICE theory too can be extended by incorporating an approach to non-technical functions in fields such as economics and administration. This would no doubt require modifications and perhaps the instrumentalization theory has indications useful to such an extension of the range of the ICE theory.

Much philosophy of technology is concerned with the human relation to technical artifacts and nature. Various models have been proposed, among which the two considered in this chapter. They have in common a focus on the technical relation as a specific type of action in the world. Both theories delve far more deeply into the nature of this relation than the standard accounts which emphasize control as the single most prominent aspect of technical action. That emphasis often leads to simplistic evaluations of technology. But before one can reach such normative conclusions, a better understanding of the significance of technology in human life is required. Further work on both approaches will contribute to that goal.

References

Böhme, G. (2012). *Invasive technification: Critical essays in the philosophy of technology*. London: Bloomsbury.

Borgmann, A. (1984). *Technology and the character of contemporary life*. Chicago: University of Chicago Press.

Cavanagh, T. (2008). Diverse designing: Sorting out function and intention in artifacts. In P. E. Vermaas, P. Kroes, A. Light, & S. A. Moore (Eds.), *Philosophy and design: From engineering to architecture* (pp. 301–315). Dordrecht: Springer.

Feenberg, A. (1977). Transition or convergence: Communism and the paradox of development. In F. Fleron Jr. (Ed.), *Technology and communist culture: The socio-cultural impact of technology under socialism* (pp. 71–114). New York: Praeger Publishers.

Feenberg, A. (1991). *Critical theory of technology*. New York: Oxford.

Feenberg, A. (1992). Subversive rationalization: Technology, power, and democracy. *Inquiry, 35*(3/4), 301–322.

Feenberg, A. (1999). *Questioning technology*. New York: Routledge.

Feenberg, A. (2005). *Heidegger and Marcuse: The catastrophe and redemption of history*. New York: Routledge.

Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). New York: Harper and Row.

Heidegger, M. (1977). *The question concerning technology* (W. Lovitt, Trans.). New York: Harper and Row.

Heidegger, M. (1995a). *The fundamental concepts of metaphysics: World, finitude, solitude* (W. McNeill & N. Walker, Trans.). Bloomington: Indiana University Press.

Heidegger, M. (1995b). *Aristotle's metaphysics 1–3: On the essence and actuality of force* (W. Brogan & P. Warnek, Trans.). Bloomington: Indiana University Press.

Houkes, W., & Vermaas, P. E. (2010). *Technical functions: On the use and design of artefacts*. Dordrecht: Springer.

Kroes, P., & Meijers, A. (2002). The dual nature of technical artifacts: Presentation of a new research programme. *Technè, 6*(2), 4.

Lukács, G. (1971). *History and class consciousness* (R. Livingstone, Trans.). Cambridge, MA: MIT Press.

Marcuse, M. (1964). *One-dimensional man*. Boston: Beacon.

Marcuse, M. (2005). In R. Wolin & J. Abromeit (Eds.), *Heideggerian Marxism*. Lincoln: University Nebraska Press.

Merleau-Ponty, M. (1955). *Les Aventures de la Dialectique*. Paris: Gallimard.

Pinch, T., & Bijker, W. (1987). The social construction of facts and artefacts. In W. Bijker, T. Hughes, & T. Pinch (Eds.), *The social construction of technological systems* (pp. 17–50). Cambridge, MA: MIT Press.

Preston, B. (1998). Why is a wing like a spoon? A pluralist theory of functions. *The Journal of Philosophy, 95*(5), 215–254.

Simondon, G. (2009). Technical mentality (A. De Boever, Trans.), *Parrhesia*, 07, 7–27.

Zimmerman, M. (1990). *Heidegger's confrontation with modernity: Technology, politics, art*. Bloomington: Indiana University Press.

Andrew Feenberg is Canada Research Chair in Philosophy of Technology in the School of Communication at Simon Fraser University, where he directs the Applied Communication and Technology Lab. He has authored or edited over ten books, including *Critical Theory of Technology* (Oxford University Press, 1991), *Questioning Technology* (Routledge, 1999*), Heidegger and Marcuse: The Catastrophe and Redemption of History* (Routledge 2005), and *Between Reason and Experience: Essays in Technology and Modernity* (MIT Press, 2010). His most recent book is *The Philosophy of Praxis* (Verso 2014).

Chapter 17
Function and Finitism: A Sociology of Knowledge Approach to Proper Technological Function

Pablo Schyfter

Abstract Philosophers of technology have dedicated considerable attention to the topic of technological functions. Many authors, including those involved in the Dual Nature of Technical Artefacts programme, have sought to produce a definition of 'proper functions' to distinguish between what technological artefact *can* do and *are meant* to do. A number of these authors are also engaged in the ongoing 'empirical turn' in the philosophy of technology, which often embraces the notion of fixed proper functions. In this article, I argue that 'proper functions' arguments are deterministic; that is, they define correct use as that which corresponds to an antecedent, fixed proper function. I present an alternative conceptualisation of technological function, based on the sociology of knowledge. Using finitism, a series of tools developed by the Edinburgh School, I posit that proper functions are socially-endorsed use. That is, technological function follows conventional social practice. Finitism can serve the 'empirical turn' because it offers analytic tools and methods to clarify the concept of technological function using empirical investigation.

Keywords Technological function • Proper function • Sociology of knowledge • Finitism

Technological function has received considerable attention from those who champion an empirical turn in the philosophy of technology.[1] In trying to clarify the concept, philosophers have presented many formulations of function, each of which gives weight to different issues. Some focus on the material properties of artefacts which enable functional action; others highlight intentionality in design and use. Many have tried to produce a definition of function centred on the causal history of design, and most subscribe to the idea of sole 'proper functions.' That is, most

---

[1] For example, consult volume 37, issue 1 of *Studies in History and Philosophy of Science*. Also consult: Houkes and Vermaas (2010), Preston (2013), and Kroes (2012).

P. Schyfter (✉)
University of Edinburgh, Edinburgh, UK
e-mail: P.Schyfter@ed.ac.uk

M. Franssen et al. (eds.), *Philosophy of Technology after the Empirical Turn*,
Philosophy of Engineering and Technology 23,
DOI 10.1007/978-3-319-33717-3_17

accounts of technological function agree that each particular kind of technological artefact is characterised by a distinct function. This ‘proper function’—what the artefact is *supposed* to do and what it is *meant* to do—defines the artefact and serves to disqualify other, ‘accidental’ uses to which the object may be put. Theories that use ‘proper function’ accept a deterministic account of technological functionality. ‘Proper functions’ are fixed at the point of design and serve to determine the correctness of all technological use to come.

Function determinism is broadly accepted, but empirically-oriented philosophies of technology are served better by theories that focus on real-world, empirical phenomena. Function finitism, a position and collection of analytic tools grounded in the sociology of knowledge, conceptualises function as a product of social practice. That is, finitism views proper function as socially-endorsed use: open-ended, indeterminate, and conventional. Crucially, finitism is “this-wordly, concrete and causal.” (Bloor 1997, p. 20) It makes sense of function by way of empirical study. Scheele writes that “the point of analysing functions is to have a tool for evaluating use” (2006, p. 29). Function finitism instead holds that the point of analysing use is to understand and explain function.

Here, I present and argue for a finitist conceptualisation of technological function. First, I outline key tenets from the ‘Dual Nature of Technical Artefacts’ programme. These include materiality, designer and user intentions, normativity, and technological function. I do not view the Dual Nature programme and the ‘empirical turn’ in the philosophy of technology as one and the same project. Nonetheless, I see many important overlaps. These include shared commitments and conceptualisations. Among these is a set of ideas and arguments concerning technological function, and especially, *proper* technological function. Second, I discuss this concept of ‘proper functions,’ and present the limitations that follow from its deterministic character. Third, I explain the tenets of meaning finitism. In simple terms, finitism views term meaning as a social institution produced and sustained by instances of term use. Fourth, I develop my finitist understanding of technological function, and contrast it with the function determinism of ‘proper functions.’ Last, I study the relationship between finitism, intentionalism, and idealism. I demonstrate that finitism is not compatible with either of the two.

An empirical turn in the philosophy of technology demands methods and analytic tools for empirical study, and ‘proper function’ offers less in this regard than other perspectives might do. Finitism, as a collection of tools from the sociology of knowledge, serves as an epistemological foundation for empirical methodology and research. It offers a way to elucidate concepts—such as function—using empirical study of real-world, concrete artefacts and users. That is, it offers the ‘empirical turn’ a remedy for its shortfall in empirical tools.

17.1 The Dual Nature of Technical Artefacts and 'proper functions'

The Dual Nature programme obtains its name from its most important claim about technological artefacts. The programme's authors argue that technological artefacts possess two fundamental characters. First, technological artefacts are physical objects: things present in space and time, with particular material qualities, capacities, and limitations. People fabricate artefacts materially, and those artefacts participate in material activities in the physical world. Second, technological artefacts are characterised by intentionality. Designers and fabricators intentionally make artefacts for something and someone, and users employ artefacts intentionally in the pursuit of particular ends. Neither of these 'natures' captures by itself what a technological artefact is; only together can they deliver a complete understanding of what make an artefact a *technological* artefact. Crucially, something must reconcile materiality and intentionality, which otherwise stand apart. The Dual Nature authors present technological function as a 'drawbridge' for this reconciliation—as something that amalgamates the two natures. Peter Kroes and Anthonie Meijers argue that technological artefacts differ from other physical constructions in one fundamental way:

> These artefacts have a purpose or function: they are objects to be used for *doing* things and are characterized by a certain 'for-ness.' (2006, p. 1)

Technological artefacts *do*, and they do *by us* and *for us*. 'For-ness' also makes manifest technological artefacts' dual nature. Function relies on the material configuration and physical capacities of technological artefacts; it also depends on people intentionally designing, making, and using those artefacts. As I noted, there exists a rift between those two natures.[2] Kroes and Meijers go on:

> If functions of technical artefacts are seen primarily as realised in the physical objects involved, the question remains how these functions are related to the mental states of human individuals… If, alternatively, functions are seen primarily as patterns of mental states, and exist, so to speak, in the heads of designers and users of technical artefacts only, it becomes mysterious how a function relates to the physical substrate of a particular artefact. (2006, p. 2)

Simply put, technological function presents a dual nature challenge. However, argue the Dual Nature authors, function also presents a solution to this quandary. Function serves as "a bridging concept that relates the physical and intentional domain" (Kroes and Meijers 2006, p. 2). As such, many philosophers of the Dual Nature programme make the study of technological function their central charge.

---

[2] The Dual Nature authors argue that this twofold nature reflects two conceptualisations of the world, physical and intentional, which must also be harmonized in the philosophical study of technological artefacts.

17.1.1 Function and Ontology

Kroes and Meijers argue that 'for-ness' distinguishes technological artefacts from other artefacts (2006). Kroes writes: "it is by virtue of its function that the object is a *technological* object" (2000, p. 28, my emphasis). For instance, people identify and employ a particular object as a 'corkscrew' because it can and is intended to remove corks from wine bottles. The corkscrew's 'for-ness' first distinguishes it from non-technological artefacts, and then differentiates the corkscrew from other technological artefacts with 'for-nesses' of their own. For Kroes, kinds of technological functions and kinds of technological artefacts are interdependent (2012). Consequently, studies of technological artefacts ought to investigate what a technological artefact *is*—its ontology—and what it is *for*—its function—jointly.[3]

Kroes argues that when something "becomes a means to an end," it "acquires a function" (2000, p. 38). Because function and ontology are interdependent, the moment at which an artefact becomes a means to an end is also the moment at which it becomes a particular kind of artefact. As such, those concerned with the nature of technological artefacts ought to ask and answer: when does an artefact become a means to an end? The Dual Nature programme holds that a satisfactory answer to this question must address both when the artefact becomes physically capable of realising its function, and when the artefact becomes embedded in some form of intentionality. Both moments, many authors of the programme contend, occur during design.

17.1.2 Proper Functions

The Dual Nature authors argue that designers define both what artefacts are and what they are for, and do so in such a way as to secure ontological and functional fixity. If artefact ontology and artefact function are interdependent, then plasticity in one implies plasticity in the other. Without a fixed function, a technology artefact lacks a rigid ontology. And yet, those who possess and employ artefacts can and do make use of them in many ways. For instance, people employ chairs to reach high shelves and cups to hold pens and pencils. Intuitively, one can describe such uses as of a different kind from what the chair and cup are 'actually' for. And yet, in both cases the materiality of the artefact successfully enables, and people intend to carry out, the rogue use. The chair is physically capable of realising what the user intends to do: materiality and intention are aligned by way of a function. One might argue that the dual nature understanding of technological artefacts supports the conclusion that for this user at this moment, the chair is something for standing. Seemingly, the artefact's ontology and function can and do change from moment to moment at the

---

[3] Elsewhere, I have developed a sociology of knowledge analysis of the Dual Nature programme's understanding of ontology and function (2009).

behest of individual users. Of course, the Dual Nature programme finds such plasticity problematic. Technological ontology and function must enjoy a certain fixity, despite users' assorted practices. That is, the 'proper functions' of specific artefact kinds—their true, justified use—must be distinct from 'accidental functions'—other, spurious uses.

The concept of 'proper function' first appeared in the philosophical study of *biological* functions,[4] but has gained traction in the study of technological ones.[5] Most simply, the proper function of a technological artefact is "*what an artefact is meant for*," and not anything that "an artefact *can* be used for" (Scheele 2006, p. 25). Crucially, proper functions must and do remain fixed despite changes in individual and collective intentions and actions. Scheele writes:

> … we may identify proper functions independently of current subjective or intersubjective intentions and are as such objective views. It is not my current personal intention, or our communities' intentions that determine the proper function, but the causal history of the artefact formation accounts for it. (2006, p. 27)

Proper functions are not contingent on the particularities of users or communities of users. Rather, designers set proper functions in advance and independently of such particularities. Once those designers specify them, proper functions do not respond to idiosyncrasies in use. Instead, use that does not conform to the proper function is incorrect. For Radder, proper functions are "those intended by designers," whereas improper functions are "those intended by other people" (2009, p. 889). More recently, Houkes and Vermaas have written:

> The role of designing is professionally played by some agents and not by others, who are typically only involved in using artefacts. This 'right to design' comes with privileges, most notably that of determining the proper use of an artefact. (2010, p. 114)

Again, designers are elevated above users in fixing proper technological functions. Though Houkes and Vermaas use terms like 'role' and 'privileges,' they do not offer an argument that acknowledges the conventional character of these phenomena. As such, the designers' standing and the proper functions they specify retain their fixity.

Technological makers specify proper functions, but ultimately users exercise those functions. Consequently, those who produce artefacts must communicate proper functions to those who employ artefacts. Only then can proper function exist as correct use. Vermaas and Houkes argue that technological artefacts are "objects that are embedded in use plans" (2010, p. 148). These plans, produced by designers, are "a series of such actions in which manipulations of [an artefact] are included as contributions to realising the given goal" (2006, p. 7). Kroes et al. similarly posit that "the designing of technical artefacts by engineers can therefore be seen as linked to the specification, or even design, of use plans" (2009, p. 567). These five authors characterise use plans as something central to the correct use of a

---

[4] See e.g. Millikan (1984, 1998, 1999a, b). Elsewhere, I have developed a sociology of knowledge analysis of philosophical theories of biological function, including 'proper functions.' (2015).

[5] Beth Preston has recently offered a criticism of this use of philosophy of biology (2013).

technological artefact for its proper function. Often, manuals and instructions communicate use plans to those still learning to employ the artefact. The authors argue that “a user who does not know the [use plan] will be clueless” (Vermaas and Houkes 2006, p. 7). ‘Clueless’ suggests both an inability to comprehend just what the artefact is and what it is for, and an inability to actualise the proper function of the artefact. In either case, designers set down what constitutes correct use. However, in more recent work, Houkes and Vermaas have written:

> Proper use is the execution of a use plan that is accepted within a certain community; improper use is the execution of a use plan that is socially disapproved. (2010, p. 93)

The pair appear to suggest a social explanation for proper functionality. However, this claim is not followed by such an argument. Rather, Houkes and Vermaas define a series of actors that can ‘justifiably’ define proper use plans. Acceptance is not social consensus, but rather privileged actors setting down correct use.

17.1.3 Normativity

The Dual Nature programme highlights the importance of technological normativity, and dedicates considerable attention to *functional* normativity. More importantly, the concept of proper functions requires an understanding of normativity. First, because proper function implies divisions between proper and accidental, ‘actual’ and other, functions. Second, proper functions also specify what the artefact *ought* to do. For instance, a corkscrew that cannot pull corks from bottles fails to accomplish its proper function; it fails at that which defines the corkscrew as a corkscrew. As such, normative evaluations concern both artefact ontology and function. Franssen writes:

> It has to be emphasised that evaluative statements like “This is a good knife” presuppose the existence of *kinds* in which artefacts can be grouped, in this case the kind “knife.” It is *as a knife* that the artefact is evaluated as good. (2009, p. 930)

Normative judgements are comparative judgements between tokens of a kind; without a plurality of examples, “it would be difficult to say in what respect the fact that the artefact was good differs from the fact that it was functional” (Franssen 2006, p. 50). Note that the dual nature conceptualisation of technological artefacts implies that those who *make* technological artefacts, and not those who *use* the artefacts, define and fix artefact kinds.

Of course, normativity also concerns use and users. It is through use that all normative evaluations of function become possible. Instances of artefact use enable assessment of artefacts’ success in fulfilling their proper function, just as they enable conclusions about users’ functional prowess. The Dual Nature programme argues that ‘use plans’ in part explain the normative character of artefact use. Franssen writes:

> … every artefact is imbedded in a use plan that specifies which operations of the artefact will lead to the end state that corresponds to the function of the artefact. A use plan tacitly or explicitly contains the circumstances that must obtain and the abilities the user must show for these operations to lead to the desired end state. (2006, p. 48)

The use plan sets down what users must do and how they ought to do it in order to realise the artefact's proper function, and in order to realise it well. The authors argue that designers set down proper functions and use plans in the making of an artefact. Thus, designers also fix the normative standards for use. What constitutes correct use is established and fixed prior to use. Proper function, established by designers, serves as an unchanging standard against which use is compared. Those who employ artefacts in accordance to proper function are realising correct use. Any divergence from the predetermined proper function results in incorrect use. Again, recent work by Houkes and Vermaas appears to grant collectives power to define other use plans (2010), but as I noted above, these arguments still give privileged actors power to define those plans. Social consensus is not given preference.

17.1.4 Proper Functions and Determinism

The Dual Nature programme argues that 'for-ness' is the most fundamental way in which technological artefacts are different from other artefacts and from each other. Function is at the heart of what makes an artefact technological. Thus, technological ontology and function are interdependent: an artefact becomes what it is when it becomes something functional. For the Dual Nature programme, designers set down and fix *is* and *for*.

For an artefact to enjoy ontological and functional stability, some proper function must define both what it is and what it is for. Proper functions—what artefacts are 'actually' for and what they are meant to do—are those specified by designers. Despite users' assorted practices, proper functions remain fixed. Any other uses are 'accidental,' illegitimate functions. Simply put: technological artefacts have specific proper functions, which are fixed before use and remain fixed despite idiosyncrasies in that use.

Technological artefacts and their functions display normative qualities. Particular artefacts can function well or poorly, and they can be good or bad tokens of an artefact kind. Importantly, users' practices can be correct or incorrect, as well as better or worse. Proper functions serve as the fixed criterion for these various types of assessment. Importantly, correct use is that which conforms to the fixed proper function.

Together, these arguments constitute a deterministic understanding of technological function. Determinism holds that design concretises proper function, and concretised proper function determines correct use. What constitutes correct use is fixed before instances of use, and correctness is concurrence with the antecedent, unchanging standard. That is, correct use follows from proper function. In contrast, finitism holds that proper function follows from socially-endorsed use.

17.2 Finitism

Meaning finitism forms part of the Edinburgh School in the sociology of knowledge. Barry Barnes and David Bloor—both prominent contributors to the School—first developed finitism as a means to understand language use, especially concept application (e.g. Barnes 1981, 1982; Barnes et al. 1996). Since then, finitism has become a more general framework with which to analyse knowledge. For instance, Bloor has used finitism to develop a sociological interpretation of Wittgenstein's arguments on rules and rule-following (1997). Martin Kusch has also contributed greatly to finitism, using it to study such epistemological topics as truth and objectivity (2002). I draw on all three scholars' work, first to present meaning finitism, and then to produce a finitist understanding of technological function.

Meaning finitism asserts that term meaning is a product of term use. Correct use of terms and concepts does not follow from an antecedent, established meaning of those terms; instead, correct use follows from how individuals in a social collective employ terms and concepts. Because meaning follows from use, meaning cannot serve to determine what constituted correct use in past cases, what constitutes it now, nor what it will constitute in cases to come. That is:

> All finitism insists upon is that there is nothing in the meaning of a term, or its previous use, or the way it has been previously defined, which will serve to fix its future proper use… (Barnes, Bloor and Henry 1996, p. 78)

Rather, correct use is an open-ended and dynamic process, whereby individual instances of term use continuously give substance to the meaning of the term. Barnes writes:

> … proper usage is developed step by step, in processes involving successions of on-the-spot judgements. Every instance of use, of proper use, of a concept must in the last analysis be accounted for separately… (1982, p. 30)

Nothing pre-determines what constitutes correct use. Each instance of term use is new and demands that users make active decisions about use. One cannot have complete certainty about correctness before use. Thus, meaning undergoes a process of *continuous creation*. It is the social collective which creates and sustains the meaning of terms. As such, meaning is a social institution, a "collective pattern of self-referring activity" (Bloor 1997, p. 33).

Meaning 'itself' is an abstraction, without any agency to compel users or determine correct use. Users are not compelled by its meaning to deploy a term in a particular way. Instead, they are compelled by fellow members of the collective.

The concrete, real-world activities of the social collective explain how terms come to be used, why and how some uses are deemed correct, and why and how terms have objectivity.

17.2.1 Training

In order to understand the key tenets of meaning finitism, it is helpful to begin with training, the process by which novice term users come to learn the correct application of a term. Consider a person leaning to use the term 'tree.' Someone skilled in using 'tree' is training a novice in the meaning and correct use of the term. The teacher makes use of ostension: she points to a particular object and tells the student that the object is called a 'tree.' She indicates important material properties of the tree, also through ostension. She notes the colour of the bark and the leaves, the shape of the trunk and the branches, and the dangling fruit. The teacher does not make use of a single exemplar. As the training process proceeds, she presents more and more trees similar to the first, but also trees that differ greatly in shape or size or colour. In all cases she explains what material properties the novice can use to identify the object and correctly apply the term. The teacher may reference trees in the world or representations of trees, but the training dynamic remains the same: she points and she names. She tests the learner by presenting examples of objects and asking what those objects are called. If the learner responds 'tree,' the teacher congratulates him; if he responds anything else, the teacher reproaches him. With each exemplar and each instance of testing, the learner gradually gains proficiency.

Three aspects of the training process are crucial. First, no matter how many examples the teacher gives to the student, the number will always be *finite*. Second, the number of future instances of term use will in effect be *infinite*; no immutable limit restricts how many uses can or will occur. Third, no two objects or instances of use will be identical. The student learns from a limited set of exemplars (the trees his teacher presents), but must apply the term by himself to future, different objects (new trees he will encounter). Without doing so, he cannot demonstrate proficiency. Because each of those instance of term use involves a new object, different from those before, each instance of term use involves a new process of observation, comparison, and decision. The same will be the case as the person meets new trees during his life. No meaning of the term 'tree' can capture the physical heterogeneity of real-world trees. When learners and users observe new things and make choices about term use, they must actively deliberate and decide.

Because training "ultimately rests on finite numbers of examples," this "renders the problem of the move to the next step ineradicable" (Bloor 1997, p. 11). Meanings, definitions, and instructions cannot determine how the user will take the next step because meanings, definitions, and instructions are generalised, and instances of use are particular. None can ensure that a term corresponds to a particular object.

17.2.2 Use

The learner has now become a skilled user. The progression from learner to knower follows from the teacher's satisfaction with the learner's ability to apply the term correctly. In the face of new cases, the learner applies the term 'tree' in a way that the teacher deems appropriate. Now a skilled user, he goes on to apply the term without direct guidance. Here the problem of moving forward presents itself:

> … moving from a finite number of examples to an open-ended, indefinitely large range of future applications… there is always going to be the problem of taking the next step, of moving from previously known cases to new cases. (Bloor 1997, p. 10)

The matter is no longer the teacher testing the student's proficiency with new cases, but instead the newly-certified skilled user employing the term on a day-to-day basis. The teacher no longer guides him; her approvals and sanctions do not compel and direct his behaviour.

The user possesses a finite set of exemplars, but he contends with an unknown and effectively infinite number of future cases. His exemplars differ from each other, and all future cases will differ from the exemplars and among themselves. Barnes writes:

> There are no clearly identical, indistinguishable particulars to cluster together… Physical objects and events are never self-evidently identical or possessed of identical essences. (1982, p. 28)

People choose how to sort real-world things into kinds, and decide if a particular object belongs within or outside a particular kind. Neither choice is pre-determined because no two objects of a kind are identical, and no kind consists of a finite number of objects. The case of term use is the same. A user must evaluate new objects and decide if they fall under a particular term. Moreover, no absolutely fixed number of such objects exists. As such, term use is open-ended, and lacks any kind of a priori certainty. In a sense, each application of a term is a 'jump into the dark.'

This argument might appear to give free rein to users, and to deny any kind of stability to terms. That is, users can do whatever they want to do without any form of constraint, and thus no term has any 'actual' meaning. This conclusion about finitism is common but mistaken. Finitism rejects logical compulsion, but it does not view term use as unconstrained. Bloor writes:

> The real sources of constraint preventing our going anywhere and everywhere, as we move from case to case, are the local circumstances impinging on us… (1997, p. 20)

General, abstract meanings cannot compel, but real-world conditions can and do. Most importantly, these conditions include *the social*:

> If an individual subordinates his inclinations to the routinely accepted mode of use of a term, it is to the practice of his fellow men that he defers, not to any set of rules or instructions for use which, as it were, come with the term… Concepts cannot themselves convey to us how they are properly to be used. (Barnes 1982, p. 29)

Meaning does not determine correct use, but the social collective prevents unrestrained idiosyncrasy. Meaning is not an individual whim, but the product of coordinated social activity.

17.2.3 Normativity

Social restraint is a normative phenomenon: a process of deeming acts correct or incorrect and of instantiating approvals and sanctions, respectively. Instances of term use are correct or incorrect because the collective evaluates individual practice and delivers a decision on correctness. "Proper usage is simply that usage communally judged to be proper," Barnes writes (1982, p. 29). Because such judgement operates on a case-by-case basis—just as do individuals' decisions about term use—correctness is never predetermined. Finitism holds that normativity is open-ended and contingent.

Just as meaning follows from the practices of use, normativity follows from collective consensus and social actors' mutual-susceptibility. The social character of normativity is essential to finitism. Only something external to the individual can define correct use. Meaning cannot do so, because meaning is a product and not a cause. Rather, social actors make normativity possible:

> Consensus makes norms objective, that is, a source of external and impersonal constraint on the individual. (Bloor 1997, p. 17)

I alone cannot determine the correct use of the term 'tree,' since I am not alone in using that term. Other people around me also employ the term, and what they do affects what I do (just as what I do affects what they do).[6]

Term-use is not a 'free-for-all' because as social actors, we are mutually-susceptible. We police each other, and our acts of policing give rise, stability, and longevity to what we consider to be proper use:

> Normative standards come from the consensus generated by a number of interacting rule followers, and it is maintained by collectively monitoring, controlling and sanctioning their individual tendencies. (Bloor 1997, p. 17)

As such, improper use is not a failure to conform to a fixed meaning, but a failure to conform one's practice to that of others successfully. Being wrong is diverging from the group.

Importantly, no two instances of evaluation or two acts of policing are identical, just as no two cases of term use are identical. As such, actors do not conform to or diverge from a fixed, unchanging standard. Social collectives continuously create normative standards, and as a result, what constitutes correct use of a term changes when local contingencies change. Because social consensus drifts, correctness

---

[6] W.V.O. Quine argues something similar when he writes, 'Each of us, as he learns his language, is a student of his neighbour's behaviour; and conversely, insofar as his tries are approved or correct, he is a subject of his neighbour's behavioural study' (1969, p. 28).

drifts. We decide in the present what is right and wrong; we did so in the past as well. ‘Correct’ is what is correct *here and now*, and correct for *this group*.

Meaning finitism delivers one overarching lesson: meaning follows use. Finitism stands opposed to meaning determinism, which holds that meanings exist independently and in advance of practice. For determinism, correct term use is like “tracing out a line that is already there” (Bloor 1997, p. 20). A fixed, antecedent meaning determines correctness before any term use occurs, and serves as an unwavering standard. The concept of ‘proper functions’ is also a deterministic one. Proper functions are fixed prior to use. They determine what will count as correct use in all cases to come. They do not react to local contingencies. For the user, the line of proper function already exists, and correct use is a matter of faithfully following that line.

17.3 Function and Finitism

A finitist understanding of function instead holds that proper technological function follows from technological use. Proper function is the product of and is sustained by interacting, mutually-susceptible people who constitute a ‘we.’ It is a social institution—a collective good—and like all social institutions, it is continuously created and ever-changing. Put otherwise, function lives in use. Technological makers may intend a particular function and they may construct the artefact capable of realising that function. That artefact’s materiality may provide the underlying capacities needed to enable its function. Nonetheless, neither design nor materiality can determine what counts as correct use. Correct use is whatever the collective deems to be correct use.

17.3.1 Training and Functional Use

As with meaning finitism, a finitist understanding of technological function begins with a look at training. A skilled technological user sets out to train a novice in the proper use of a technological artefact, such as a corkscrew. She presents the artefact, highlights its key material components and their behaviour, and then proceeds to demonstrate how the artefact can pull corks from bottles’ necks. She places the tip of the artefact on the cork and begins to twist the screw. In doing so, she notes how to arrange the artefact, she points to where she placed the tip, how she grasped what part of the artefact, and how she works her hand to work the twisting. Again, the trainer uses ostension to instruct; she points and explains. She continues by noting when to stop the twisting, how to modify the physical configuration of the corkscrew, and how to work the newly-configured artefact. She points to the motion of the cork as it pulls away from the bottle, and demonstrates how she achieves her end.

Other occasions of training follow this first one, and each affords the novice a new exemplar of correct use. Eventually, the novice attempts to work the corkscrew alone, while the teacher observes his performance. She expresses approval and enacts sanctions in response to correct and incorrect use, respectively.

Training in artefact use shares with training in term use a number of crucial characteristics. First, the number of exemplars of correct corkscrew use is *finite*. The teacher demonstrates the relevant functional acts a finite number of times and supervises the learner in a finite number of cases. She also uses a finite number of artefacts in the process (perhaps only a single one). Second, future instances of functionality are in effect *infinite*. Once the novice becomes certified as a skilled user, he proceeds to new occasions of artefact use, and there exists no immutable limit to the number of those occasions. Third, no two cases of artefact use were, are, or will be identical; each occasion differs in a variety of ways. The skilled user may employ different corkscrews in different instances, but even if he uses only one token, that object does not remain static, as all material things change with time. Each instance of artefact use will involve a new, different cork. The user's hands and fingers will move in broadly similar, but never identical ways. And of course, the physics of one instance will always differ from the physics of another. For example, the following will differ: friction between cork and bottle; pressures inside and outside the bottle; temperatures of the bottle, the cork, and the surrounding environment.

The number of exemplars the teacher gives the novice is finite, and each exemplar is different. The next case will never be an exact repetition of previous ones, so new functional acts do not enjoy pre-determination based on previous occasions of use. Moreover, no general description of the artefact's proper functions, such as 'a corkscrew is for pulling corks from bottles,' can capture the heterogeneity of real-world use. Similarly, no set of instructions can determine particular cases of use. Both descriptions of proper function and instructions for use are generalised, whereas instances of use are particular. As with term use, there is a problem of moving to the next case. Put otherwise, artefact use is always open-ended; the user always 'jumps into the dark' when carrying out a technological function.

17.3.2 Proper Functions and Continuing Creation

Technological makers—those involved in design and fabrication—deliver artefacts to the world. They also present what those artefact are 'actually' meant for: their proper functions. Nonetheless, functionality lives in practice; it persists because users employ artefacts as means to ends. The concept of any given artefact's proper function is an abstraction of particular real-world acts involving particular artefacts and particular users. To understand functionality, one must look to those acts. Doing so reveals that social practices continuously create proper function.

Those who argue that proper functions determine correct use do not disregard users and their behaviour entirely. However, many view the role of users as something so obvious as to be only marginally significant. De Ridder acknowledges:

> … the trivial point that artefacts do not *do* anything without human agency. They 'work' only when we use them and as a result, an explanation of their working must include information about human action. (2006, p. 82)

Radder, though more concerned with the role of users, similarly writes, "technologies, if they are expected to keep functioning, cannot be left to themselves" (2008, p. 54). In both cases, the authors portray users as somehow in the service of artefacts. They are relevant only insofar as they actuate artefacts. As a result, analyses of technological function cannot excise users, but their role is one subservient to design and to proper function.

Finitism dedicates much more attention to users and their practices. While the necessity of users may be a trivial point, what users' actions accomplish is not. Every time that a person uses a corkscrew to open a bottle of wine, the event is a new display of what the artefact does. Together, the many examples of different users employing different corkscrews to do roughly the same thing constitute the active functioning of the artefact kind. Stated simply, corkscrews pull corks from bottles because people use corkscrews to pull corks from bottles. To say that artefacts have functions—that functions somehow exist 'inside' or are 'built into' artefacts—is a roundabout and limited way of noting that artefacts are used functionally.

Because use is a type of human practice, it is necessarily varied and always indeterminate before it occurs. That is, artefact use is open-ended. The skilled user employs previous instances of use as points of reference, but each new case will be distinct. Making use of a corkscrew to draw corks from bottles never 'looks the same.' Thus, a statement like 'corkscrews function to pull corks from bottles' is a conceptual abstraction of heterogeneous, real-world events. The generalised definition of function stands in for a messy totality of assorted enactments of function—the empirical reality of function.

Each of these varied acts of function is the doing of users. Importantly, those users do not carry out their actions in complete isolation. Other people often witness my use of corkscrews, just as I commonly witness others using similar artefacts to do similar things. Each of these instances gives to people a new demonstration of corkscrews' function, and so we are continuously presented with reminders that corkscrews are for pulling corks from bottles. Together, these instances of functionality form a body of function exemplars, from which we draw guidance and into which we deliver new examples that serve to guide others. Successfully 'taking the next step'—completing a new instance of correct artefact use—occurs in a community of others trying to do the same thing. The totality of functional occurrences rests on the acts of social agents, not isolated individuals.

As it did for term use, this finitist description seems to give free rein to users and appears to strip function of all substance and stability. Proper function constraints such as design history seemingly play no role. The finitist account also seems to ignore materiality, one of the Dual Nature programme's two natures and something that proper function advocates routinely discuss. Houkes and Meijers write:

> Engineering is not based on *anything goes*. You cannot make a hammer from foam, nor can you use foam as a hammer. (2006, p. 123)

Clearly, the case of artefact use is similar. I cannot employ a corkscrew to make a telephone call, and a foam corkscrew cannot help me to open a bottle of wine. Physical constraints without question set important limitations on what users and artefacts can do, but materiality by itself cannot determine what counts as proper artefact use. Put otherwise, the material that constitutes the corkscrew does not dictate what a user ought to do, nor if one instance of practice is proper functionality and another is not. 'Stuff' enables and constrains, but it does not determine proper function.

17.3.3 Functional Normativity and Collective Consensus

Many involved in the 'empirical turn' in the philosophy of technology have studied technological normativity, and the concept of proper functions requires an understanding of normativity. Proper functions fix what counts as correct use and serve to distinguish between good and bad tokens of an artefact kind. Because proper functions are 'built into' artefacts during the process of design, judgements of correct and incorrect hinge on designers' ambitions in producing those artefacts. Although individuals may and do employ artefacts to carry out a host of tasks, only what conforms to the design history of the artefact 'counts' as correct use. Kroes, Franssen and Bucciarelli write:

> A particular technical artefact may seem quite fit for a job that a user has in mind for it, but that in itself does not make it a rational product, since for it to be that it should also have been designed for the job. (2009, p. 574)

Proper functions are produced and fixed by design, even if users' actions include many functional acts not envisioned by designers. For finitism, all functional acts—envisioned and otherwise—matter.

Finitism understands proper function as what the social collective takes to be proper function, and correct use as what the social collective takes to be correct use. Normative judgements always occur on a case-by-case basis. Consider first judgements of users' actions. People judge each instance of artefact use when it occurs and as a distinct set of actions. Whenever I employ a corkscrew in the company of others, my actions and the artefact are susceptible to approval or admonishment, and open to correction. When those around me think my actions inappropriate or unskilful, their rebukes serve to delimit what forms of use are correct. Normativity is an ongoing, dynamic process. Importantly, judgements of use are comparative. Previous instances of use serve as the standard against which member of the collective judge new cases of use. A case of use is correct if the collective decides that it resembles previous correct cases: if the collective decides that it fits the history of correct use. As a result, the normative process is human and contingent. Correctness is not concurrence with an abstract proper function, and it is never predetermined.

Members of the social collective compare new instances of artefact use to an existing set of exemplars of correct use. If a new case is deemed correct, it enters the set of exemplars; it becomes one part of the standard people use to evaluate new cases of artefact use. Because no two instances of use are ever identical, each new example of proper use modifies the set of exemplars. As such, new use is evaluated not only in relation to previous use, but is also evaluated in relation to a standard that is constantly changing.

Consider now normative judgements of artefacts 'themselves.' That is, evaluations of the quality of a token artefact's functional performance. One corkscrew may pull corks better than does another, or do so worse in one instance than in another, or fail to do so entirely. As with the first case of normativity, the issue is functional exemplars and comparative evaluation. Once an artefact encountered is classified as a token of an artefact kind, it becomes part of an artefact exemplar set. Moreover, every occasion of functionality is made part of a functionality exemplar set. That is, every instance of a corkscrew successfully pulling a cork from a bottle is part of a set of exemplars of corkscrew functionality. What constitutes a good functional performance by an artefact is based on that set of exemplars; it is based on comparison. What we think a particular token artefact *can* do and *ought* to do rests on earlier, real-world experiences, no two of which are identical. As such, normative judgements of an artefact's functional performance involve active comparison of contingent particulars, not an abstract statement about proper function. Labelling a token a poor corkscrew, or saying that it 'doesn't work,' is a rebuke for failure to fit a set of exemplars.

Sets of exemplars—compilations of acts deemed correct and artefacts deemed good—belong to the social collective. The aggregate of users, drawing on the sum of their activities and judgements, arrives at collectively-constituted and shared consensuses. These serve as external and objective checks on individual idiosyncrasy. I can use a corkscrew in countless ways, but only certain functions are proper and only certain uses are correct. A corkscrew can do many things in many ways, but only some are ones it *ought* to do to be a good, functional corkscrew. Uses and artefacts are proper and good because the social collective to which I belong deems them so. Consensus about use enjoys stability because individuals police each other's particularities, and because each of those individuals is susceptible to criticism and correction. Consensus about artefacts enjoys stability because individuals police and judge artefacts. However:

> Finitism doesn't imply that if you examine individual thoughts you will find meaning is indeterminate, but if you bring in the community this indeterminacy is removed or corrected. It can never be removed. Consensus may furnish us with norms, but it does not overcome finitism. (Bloor 1997, p. 26)

Finitism does not replace the intractable fixity of proper function with a similar intractable fixity of social consensus. Consensus enjoys, as do all social institutions, stability, but it will never be immutable. No two exemplars of proper function or good artefact are identical, and each case of normative evaluation adds a different exemplar to the relevant set. As a consequence, consensus drifts over time.

For determinism, proper function lays down a path of correct use in advance of users' practices. Correct use is the loyal following of that path, and incorrect use is a failure to do so. As such, one needn't explain correct use beyond noting the proper function of an artefact, which is fixed during the process of design. Incorrect uses are the product of localised influences that cause users to err in their actions, to "deviate because of certain social facts" (Scheele 2006, p. 33). Deterministic accounts portray fidelity to the invariant paths of proper function as correct, rational behaviour. Divergence implies lack of reason:

> Society, or certain communities, may simply be irrational and, for all kinds of social reasons, decide otherwise about the proper use of an artefact. (Scheele 2006, p. 35)

For determinism, there exists one single, fixed proper function to which erring persons can be returned.[7]

In contrast, finitism holds that cases of correct and incorrect use are both explicable with reference to the acts of users and the collectives of which they are members. As I noted, whether artefacts accomplish their proper functionality is also a conventional judgement of the collective. There exists no proper functionality that fixes a correct path in advance to use. Instead, proper function and correct use are the constant cutting of a path by those making use of technological artefacts. Improper functions and incorrect use are the result of leaving the group cutting that path.

17.3.4 Functional Use and Finitism

In my summary of the concept of 'proper functions,' I argued that it forms part of a deterministic conceptualisation of technological function. Technological makers define and fix what a particular technological artefact is 'actually' for: its proper function. Only those uses which conform to this proper function count as correct uses. As such, makers concretise what counts are correct use before any use at all. That is, there exists a deterministic relationship between proper functions and correct use.

A finitist understanding of function views proper function not as something concretised and antecedent to use. Instead, finitism argues that proper function is socially-endorsed use. Social collectives continuously create proper functions with

---

[7] Beth Preston makes a comparable point in writing that "there is something specific [technological artefacts] are supposed to do, even though they may never perform this function, or may be temporarily coopted for some other use" (1998, p. 215). However, Preston does not appear to subscribe fully to a deterministic account, as she also writes that "we have come to understand [action and function] as constructed through the constant interaction of individuals with their environment and with each other" (2013, p. 187). The second point shares a great deal with finitism.

each instance of use that the collective deems correct. The collective deems use correct or incorrect based on comparisons with assorted exemplars of functionality. Stated simply, proper function is a convention.

17.4 Finitism, Intentionalism and Idealism

Philosophical studies of technological function have taken great interest in the issue of intentions. For example, intentions have served to distinguish biological and technological functions. Intentions also form a central part of theories that draw attention to causal histories in technological design and fabrication. Importantly, authors have examined and theorised intentions in many different ways in order to make sense of proper functions and correct use. Last, some authors subscribe to intentionalist theories of function. Preston wrote that intentionalism "takes artefact functions to depend entirely on the intentional states of human agents" (2009, p. 225). That is, individual and/or collective intentions alone determine technological functions. Both intentionalism and finitism focus on human agents, but finitism focuses on *practice*, not intentions. Though some authors synonymize the two terms, there exist important differences.

Beth Preston and other philosophers of technology commonly present J.R. Searle's work on social reality as archetypal intentionalism. Searle argues that all functions—natural and technological—are 'observer relative.' Functions are not intrinsic properties of things, but rather characteristics that rest on the 'assignment of function.' Searle describes this 'assignment of function' as a "feature of intentionality" (1995, p. 14). As such, function rests strictly on intentions. More specifically, function rests on *collective* intentionality. Collectives of people have a collective intentionality not because they "engage in cooperative behaviour," but because they "share intentional states such as beliefs, desires, and intentions" (Searle 1995, p. 23). Searle argues that this sharing is not an abstraction of similar but distinct individual intentions. Rather, the intention is one possessed by the group itself.

Finitism does not subscribe to this position. Finitist studies of term use (Barnes 1982; Barnes et al. 1996) and rule-following (Bloor 1997) explicitly reject any position that reduces collective goods to individual intentions or dispositions. My own finitist understanding of technological function likewise rejects any theory that view functions as explicable solely by intentions. For finitism, the individual serves a role only as an accepted member of a social collective. Moreover, the role hinges not on mental phenomena like intention, but on socially-situated practices like speech acts and physical actions. The function of a corkscrew is to pull corks from bottles not because the collective intends to use corkscrews in this manner. Instead, the artefact has this function because interacting and mutually-susceptible people make reference to corkscrews as objects for pulling corks from bottles, and because the same people use corkscrew to pull corks from bottles. Intentions play no role in the finitist explanation of how proper functions come to be and what counts as correct use. The focus is practice. Searle makes sporadic references to 'doing,' which at first appear to be mentions of practice. However, he presents 'wanting' and 'believing' as examples of doing. Like his theory more broadly, Searle's understanding of what people *do* rests on an understanding of intentions, and not practices.

Finitism's focus on practice also counters criticisms of idealism. Philosophers of science and epistemologists have criticised the sociology of knowledge, and the Edinburgh School in particular, for what they perceive to be 'anti-realism'. Bloor summarises the criticism:

> … sociologists of knowledge portray the world as if it depended on belief, rather than belief depending on how things stand in the world. In other words, the accusation is one of *idealism*. (1996, p. 839)

I believe that a finitist understanding of technological function is likely to draw similar criticism, particularly as recent philosophy of technology draws great attention to the material qualities of technological artefacts and behaviours. Intentionalist theories, including Searle's, receive criticism for what some view as an overlooking of these material qualities. Properties such as the metal that makes up the corkscrew and the force it exerts when being used help constitute, at least in part, the artefact's function. I agree that a satisfactory theory of technological function cannot disregard the physical, and finitism does nothing of the sort. Finitism focuses on the practices of social agents, and those practices are real-world, material phenomena. My use of a corkscrew occurs in space and time. It involves a physical entity. It is enabled by the artefact's material qualities, and circumscribed by real-world material limitations. Ultimately, one cannot understand this type of social practice without making sense of materiality. In this sense, a finitist explanation of technological function and correct use incorporates a greater concern for materiality than does the more abstract theory of 'proper functions.'

17.5 Conclusion

A finitist understanding of technological function rests on three intertwined claims. First, there exists no 'proper function' that can fix future correct artefact use. Put otherwise, proper function is not a "specification, or template, or algorithm fully formed in the present" (Barnes et al. 1996, p. 55) that can set down what will count as correct use in cases to come. Second, technological function is a social institution, continuously created by interacting and mutually-susceptible social agents. What counts as correct artefact use is decided on a case-by-case basis by the collective. Third, no use of an artefact is ever indefeasibly correct. Instances of artefact use must be compared to exemplars of correct use, and the result of that comparison is never predetermined.

Finitism holds that *proper function is socially-endorsed use*. Many existing philosophical theories of proper function—such as that advanced by the Dual Nature programme—are instead *deterministic*. That is, they present proper function as a characteristic embedded in technological artefacts through the process of design.

Once established, the property is fixed. Once fixed, it determines correct use now and in cases to come. Thus unlike function finitism, function determinism divorces proper function from the particularities of real-world practice and social collectives.

Ultimately, an empirically-oriented philosophy of technology is better served by finitism, because the perspective argues that a function never actuated is no more than an abstraction. Function is real-world practice, and thus demands empirical study. Such research forms a useful link between conceptual discussions of function and the empirical realities of people using functional artefacts to meet specific ends. Recent philosophy of technology has traced a different path from earlier, deterministic philosophies of technology (Kroes and Meijers 2000). Finitism—"this-wordly, concrete and causal" (Bloor 1997, p. 20)—can accomplish something similar by undermining deterministic theories of function, and by enabling a more dedicated turn to the empirical.

References

Barnes, B. (1981). On the conventional character of knowledge and cognition. *Philosophy of the Social Sciences, 11*, 303–333.

Barnes, B. (1982). *T.S. Kuhn and social science*. London: Palgrave Macmillan.

Barnes, B., Bloor, D., & Henry, J. (1996). *Scientific knowledge: A sociological analysis*. London: Athlone.

Bloor, D. (1996). Idealism and the sociology of knowledge. *Social Studies of Science, 26*(4), 839–856.

Bloor, D. (1997). *Wittgenstein, rules and institutions*. London: Routledge.

de Ridder, J. (2006). Mechanistic artefact explanation. *Studies in History and Philosophy of Science, 37*(1), 81–96.

Franssen, M. (2006). The normativity of artefacts. *Studies in History and Philosophy of Science, 37*(1), 42–57.

Franssen, M. (2009). Artefacts and normativity. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 923–952). Amsterdam: Elsevier.

Houkes, W. (2006). Knowledge of artefact functions. *Studies in History and Philosophy of Science, 37*(1), 102–113.

Houkes, W., & Meijers, A. (2006). The ontology of artefacts, the hard problem. *Studies in History and Philosophy of Science, 37*(1), 118–131.

Houkes, W., & Vermaas, P. (2010). *Technical functions: On the use and design of artefacts*. London: Springer.

Kroes, P. (2000). Engineering design and the empirical turn in the philosophy of technology. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (pp. 19–43). Amsterdam: JAI.

Kroes, P. (2012). *Technical artefacts: Creations of mind and matter*. London: Springer.

Kroes, P., & Meijers, A. (2000). Introduction. In P. Kroes & A. Meijers (Eds.), *The empirical turn in the philosophy of technology* (pp. xvii–xxxv). Amsterdam: JAI.

Kroes, P., & Meijers, A. (2006). The dual nature of technical artefacts. *Studies in History and Philosophy of Science, 37*(1), 1–4.

Kroes, P., Franssen, M., & Bucciarelli, L. (2009). Rationality in design. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 565–600). Amsterdam: Elsevier.

Kusch, M. (2002). *Knowledge by agreement*. Oxford: Oxford University Press.

Millikan, R. G. (1984). *Language, thought and other biological categories*. Cambridge, MA: MIT Press.

Millikan, R. G. (1998). In defense of proper functions. In C. Allen, M. Bekoff, & G. Lauder (Eds.), *Nature's purposes* (pp. 295–312). Cambridge: MIT Press.

Millikan, R. G. (1999a). Proper functions. In D. J. Buller (Ed.), *Function, selection, and design* (pp. 85–96). Albany: SUNY Press.

Millikan, R. G. (1999b). An ambiguity in the notion 'function'. In D. J. Buller (Ed.), *Function, selection, and design* (pp. 115–122). Albany: SUNY Press.

Preston, B. (1998). Why is a wing like a spoon? A pluralist theory of function. *The Journal of Philosophy, 95*(5), 215–254.

Preston, B. (2009). Philosophical theories of artefact function. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 213–233). Amsterdam: Elsevier.

Preston, B. (2013). *A philosophy of material culture: Action, function, and mind*. London: Routledge.

Quine, W. V. O. (1969). *Ontological relativity and other essays*. New York: Columbia University Press.

Radder, H. (2008). Critical philosophy of technology: The basic issues. *Social Epistemology, 22*(1), 51–70.

Radder, H. (2009). Why technologies are inherently normative. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 887–921). Amsterdam: Elsevier.

Scheele, M. (2006). Function and use of technical artefacts: Social conditions of function ascription. *Studies in History and Philosophy of Science, 37*(1), 23–36.

Schyfter, P. (2009). The bootstrapped artefact: A collectivist account of technological ontology, functions, and normativity. *Studies in History and Philosophy of Science, 40*(1), 102–111.

Schyfter, P. (2015). Function by agreement. *Social Epistemology, 29*(2), 185–206.

Searle, J. R. (1995). *The construction of social reality*. London: Penguin.

Vermaas, P. E., & Houkes, W. (2006). Technological functions: A drawbridge between the intentional and structural natures of technical artefacts. *Studies in History and Philosophy of Science, 37*(1), 5–18.

Pablo Schyfter B.A. Stanford University, MSc. and Ph.D. in Science and technology studies from The University of Edinburgh. Lecturer in Science, Technology and Innovation Studies, The University of Edinburgh. Major interests: the sociology of engineering knowledge; social studies of synthetic biology; feminist science and technology studies. Current projects: Co-author, *Engineering Life* (in progress).