Duane F. Shell
David W. Brooks
Guy Trainin
Kathleen M. Wilson
Douglas F. Kauffman
Lynne M. Herr

The Unifie
Learning M

The Unified Learning Model

Duane F. Shell · David W. Brooks · Guy Trainin ·
Kathleen M. Wilson · Douglas F. Kauffman ·
Lynne M. Herr

The Unified Learning Model

How Motivational, Cognitive, and Neurobiological Sciences Inform Best Teaching Practices

Springer

Prof. Duane F. Shell
University of Nebraska-Lincoln
College of Education & Human Sciences
114 Teachers College
Lincoln NE 68588-0345
USA
dshell2@unlnotes.unl.edu

Prof. David W. Brooks
University of Nebraska-Lincoln
College of Education & Human Sciences
118 Henzlik Hall
Lincoln NE 68588-0355
USA
dbrooks@unlserve.unl.edu

Prof. Guy Trainin
University of Nebraska-Lincoln
College of Education & Human Sciences
118 Henzlik Hall
Lincoln NE 68588-0355
USA
gtrainin2@unl.edu

Prof. Kathleen M. Wilson
University of Nebraska-Lincoln
College of Education & Human Sciences
118 Henzlik Hall
Lincoln NE 68588-0355
USA
kwilson3@unl.edu

Prof. Douglas F. Kauffman
University of Nebraska-Lincoln
College of Education & Human Sciences
114 Teachers College
Lincoln NE 68588-0345
USA
dkauffman2@unl.edu

Dr. Lynne M. Herr
Educational Service Unit 6 (ESU 6)
210 5th Street
Milford NE 68405
USA
lherr@esu6.org

ISBN 978-90-481-3214-0 e-ISBN 978-90-481-3215-7
DOI 10.1007/978-90-481-3215-7
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009937400

Printed on acid-free paper

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

Preface

This is a book about how humans learn. Our focus is on classroom learning although the principles are, as the name of this book indicates, universal. We are concerned with learning from pre-school to post-graduate. We are concerned with most business, industrial and military training. We do not address how infants learn how to speak or walk, or how grown-ups improve their tennis swing. We do address all learning described by the word "thought", as well as anything we might try to teach, or instruct in formal educational settings.

In education, the words *theory* and *model* imply conjecture. In science, these same words imply something that is a testable explanation of phenomena able to predict outcomes of experiments. This book presents a model of learning that the authors offer in the sense of scientists rather than educators. Conjecture implies that information is incomplete, and so it surely is with human learning. On the other hand, we assert that more than enough is known to sustain a "scientific" model of learning.

This book is not a review of the literature. Instead, it is a *synthesis*. Scholars and many teachers likely have heard much if not most or even all of the information we use to develop the unified learning model. What you have not read before is a model putting the information together in just this way; this is the *first* one.

We do indeed pick and choose from the available knowledge to create this synthesis. What we do not do is overlook certain facts or data, or shape the data to fit our model. To the best of our knowledge, we are able to account for *all* of the known *data* about learning.

We do not necessarily account for anecdotal information. For example, there are many legends regarding autistic savants. If savants really do spontaneously show skills in the absence of learning activities, then our model is wrong. There are abundant anecdotal reports of such savant skills. When studied closely, however, savants appear to learn in the same way as other humans. Autistic savants seem to be flawed in not being able to learn as broadly as most of us can. We speculate that these savants become narrowly skilled not because of special gifts but only because those narrow skills are the ones that they can most easily attain.

Why did we write a book? Our goal is to reach those entrusted to guide the learning of other people. Teachers are an understandably skeptical audience. If the truth be told, teachers have been trained in so many ideas (fads) that controvert their experience that they don't really believe much that is labeled with the term theory. This model has something very important going for it, however. It explains and predicts how learners in classrooms actually do behave and learn. Teachers will see that right away. What's very important is that, in a very critical and fundamental manner, the model informs us about ways we can try to help our students learn better.

It's one thing to tell a story that classroom teachers can believe; it's quite another to tell one that scholars and researchers of teaching and learning will find acceptable. That left us with a serious challenge; could we create one book that addressed both groups? We decided that, rather than write two books or a series of papers, we would write a book in which the chapters had extensive, detailed notes and comments. The notes are aimed at researchers, and include citations and arguments used to justify what appears in the text.

This book has six authors. Duane Shell is an educational cognitive psychologist. His knowledge provided the background from which most of the rest of the model was developed. David Brooks is a chemist who once studied the mechanisms of enzyme action. That work involves figuring out the details of atoms and molecules that one never sees; he is trained in developing models about how things that he can't see actually work. It was Brooks' insight accounting for motivation that opened to path to the unified learning model. Guy Trainin and Kathy Wilson both are educational psychologists who prepare classroom teachers and study strategies for effective teaching. They checked the entire manuscript to ensure that no claims about classroom teaching were made that did not enjoy research support. Doug Kauffman is an educational psychologist who studies how feedback and motivation impact students' academic engagement and achievement. Lynne Herr is both an instructional technology consultant and a technical writer. She worked to make a book in which very complex technical content is co-mingled with more ordinary content understandable to those well-trained teachers for whom cognitive science is new.

Contents

Chapter 1
The Unified Learning Model

The Unified Learning Model (ULM) is a model of how people learn and a resulting model of teaching and instruction. The academic literature is filled with models about learning, teaching and instruction. The most obvious question then is, "Why do we need another theory/model of learning?" Our answer is that the current literature contains only limited theories about isolated specific learning and instructional phenomena. As a result each of these theories explains some, but not all learning phenomena. In addition, each tends to have its own vocabulary. The result is a hodge-podge of specific learning principles and teaching guidelines that often seem in conflict with each other.

The ULM, as its name suggests, is a unifying synthesis of these existing theories. The ULM is not based on some revolutionary new research findings on how people learn. In fact, while recent advances in neurobiology and brain science have enlightened our understanding of the underlying neural mechanisms involved, we have known how people learn for a long time. There are very mature and well researched areas within the broad field of learning and teaching. What the ULM does is bring these disparate topic areas together under a single umbrella. It connects them together with simplicity and clarifies the ways in which they are interconnected.

The ULM accomplishes this by focusing on the basic processes and components of learning. Our contention in this book is that the components of the ULM underlie all learning phenomena. Hence, all current models and theories of learning, teaching, and instruction can be subsumed within the ULM. Our goal with the ULM is to replace the current diverse and confusing array of learning concepts and terminology with a scientifically grounded concise set of core learning principles. If you understand these ULM principles, you understand how learning occurs and how this learning can be facilitated by teaching and instruction.

Following what is commonly referred to as Occam's Razor (or the rule of parsimony), we aim to suggest a model that is simple while explaining all observed phenomena.[1] So what are the components that underlie all learning phenomena? In the ULM there are three: working memory, knowledge, and motivation.

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_1,

Working Memory

The centerpiece of the ULM is working memory. Working memory is where temporary storage and processing of information happen in the brain. Suppose someone were to read to you a series of single digit numbers at a rate of one per second (for example, "one, three, seven, four . . ."). You then are expected to recite back those same digits. Most of us can recall around seven digits without error. With just a little practice, we can do much better – say 15–20 digits. The number of digits is a crude measure of your working memory capacity, and "the place" where you do this is called your working memory. Working memory is central to all current models of cognition and neurobiology. In the scientific literature, one cannot talk about thinking, attention, decision making, brain functioning, or, most importantly, learning without talking about working memory. Understanding working memory is the key to understanding learning. So you may ask, "If working memory is so important, why have I never heard much about it?" "Why isn't working memory the primary topic in every pre-service or in-service education course?" We have asked these questions ourselves. Our answer is this book.

The way working memory functions dictates how learning happens and what instructional methods and techniques facilitate or hinder learning. A science of learning, teaching, and instruction must be based to a great degree on the science of working memory. The ULM is this working-memory-based science of learning. We will spend considerable time discussing how working memory operates, how working memory produces learning, and how the operation of working memory can be influenced through teaching.

Knowledge

The ULM, however, includes more than working memory. The second core component of the ULM is knowledge. In the scientific literature of cognitive psychology, cognitive science and neuroscience, knowledge means something very different than the way educators typically think of the term. Educators usually think of facts and general concepts when they hear "knowledge." For example, think of the first level of the original Bloom's Taxonomy, the "knowledge" level. A revision of the original taxonomy calls this the "remember" level.[2] In the scientific literature, however, knowledge means *everything* that we know. It not only means facts and concepts, but also problem solving skills, motor behaviors, and thinking processes. Every category of Bloom's Taxonomy, then, is knowledge to a cognitivist. Knowledge is kept (or stored) in long-term memory. Psychologists generally just call this *memory* and drop the "long-term" modifier. Memory is the cognitive term for the brain and nervous system, thus knowledge is everything we know or can do that is stored in our memory.

Knowledge has a two-fold role in the ULM. First, knowledge is the goal of the ULM. The purpose of learning is to increase the many facets of our knowledge.

Learning has occurred when our store of knowledge is increased or changed. In one very important sense, knowledge is the outcome or result of the operation of working memory. It is working memory's product. Knowledge, however, has a second function. Knowledge influences how working memory operates. The things working memory can do are affected by existing knowledge. In relation to learning, you may have heard of this as "the prior knowledge effect": the more you know about something, the easier it is to learn something new about it. So knowledge is a process of working memory as well as its product. We will devote much time to discussing how knowledge is increased, how it becomes more sophisticated, how it moves through Bloom's hierarchical taxonomy, and how this increasing knowledge impacts future learning.

Motivation

The third and final component of the ULM is motivation. Educators are immediately aware of how important student motivation is in the classroom. Motivation is discussed often in education. It is framed in terms of things like interest and preference or building students' self-confidence and self-esteem or rewards or goals. Motivational ideas and constructs are seemingly everywhere, and motivation is currently one of the most highly researched topics in education.

The ULM has a very specific role for motivation. Motivation is the impetus for directing working memory to a task; in our case, directing working memory to the task of learning. To our knowledge, the ULM is the only model of learning or motivation that explicitly links motivation to working memory.[3] Motivation is an inherent component of working memory operation and plays a critical role in effective and efficient allocation of working memory to learning. Understanding how motivation works in conjunction with working memory will help teachers understand how the various motivational constructs they have heard about actually work to motivate students to learn. So we will spend much time examining how motivation and working memory operate.

Three Principles of Learning

The ULM is founded on three basic principles of learning:

1. Learning is a product of working memory allocation.
2. Working memory's capacity for allocation is affected by prior knowledge.
3. Working memory allocation is directed by motivation.

These three principles of learning form the foundation for a complete theory of instruction and teaching. Simply put, teaching that follows these principles will be effective; teaching that does not follow these principles will be ineffective.

The remainder of this book is divided into two sections. The first will cover the three basic components of the ULM to explain how working memory, knowledge, and motivation work to produce learning. This is the underlying model of learning that forms the foundation of the ULM and from which the three basic principles are derived. The second will use the three basic principles of learning to develop recommendations for successful teaching and instruction.

Notes

1. http://en.wikipedia.org/wiki/Occam%27s_razor (Accessed March 22, 2009).
2. The original taxonomy is found at: Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: D. McKay. The revised taxonomy is found at: Anderson, L., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom′s taxonomy of educational objectives*. Columbus: Merrill. The original levels were labeled knowledge, comprehension, application, analysis, synthesis, and evaluation. The labels in the revised taxonomy are remember, understand, apply, analyze, evaluate, and create.
3. Hayes offered a framework that included motivation and working memory together as part of a scheme for understanding writing. The ULM is a general learning model that applies to all learning (not just how to write) and makes specific the role of motivation in the learning process. Hayes, J. (2000). A new framework for understanding cognition and affect in writing. In R. Indrisano & S. J. Squire (Eds.), *Perspectives on writing: Research, theory, and practice* (pp. 6–44). Newark, DE: International Reading Association.

Part I
Developing the Unified Learning Model

Chapter 2
Learning

Understanding how neurons work together to generate thoughts and behavior remains one of the most difficult open problems in all of biology, largely because scientists generally cannot see whole neural circuits in action.

Meisenböck[1]

What is learning? A seemingly simple question, but the answer is both simple and complex. At its most basic, learning is a relatively permanent change in a neuron. Over the past half-century, scientists have uncovered most of the basic biological and chemical processes involved in learning at the neural level. Learning results when the synaptic potential of a neuron, the likelihood of a neuron transmitting an electric potential, is systematically changed.[2]

As you might expect, the biological and chemical processes involved in this change are extremely complex. Luckily, in the same way that an engineer building a bridge doesn't need to deal with the subatomic structure of the atoms that make up the materials she is building with, educators do not need to deal with the underlying biochemistry of the neuron to develop effective teaching and instruction. The ULM, however, derives its principles of learning from the neurobiology of learning. So this is where we start.

The Neurobiology of Learning

The brain and nervous system exist in large part to take in information from the world and use that information to direct motor action in the world. Although the brain and peripheral nerves do perform other basic and often automatic functions such as control of physical, hormonal, and body regulatory functions, like heart beat, we are concerned with those parts of the brain and nervous system involved in higher learning and behavior, that is, the parts that matter for school learning.

These parts of the brain are situated primarily in the cortex. Once we move beyond the brain areas involved in basic biological functioning, the vast majority of the remainder of the brain is devoted to gathering sensory inputs and generating motor actions. The brain has two primary jobs. One is to take in and save information about the world from the senses (hearing, taste, touch, smell, vision, and

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_2,

internal feedback from those internal sensors that sense how things are "working together" called proprioceptive feedback).[3] The other is to produce motor outputs that generate functional behaviors in the world like finding food, building shelter, and speaking. Most of the higher brain areas of the cortex are devoted to sensory input areas (the occipital lobe for vision; the olfactory cortex for smell, etc.) and motor control areas (including one-to-one mapping of physical areas of the body like fingers on the motor cortex itself). The remainder of the cortex is devoted to specialized functioning like language or is available as a general memory area.

Much of early infancy is devoted to the development of sensory and motor areas for things like learning to interpret visual input and learning basic motor skills like crawling and walking. The developmental biologist and psychologist Jean Piaget extensively studied this period of what he called the *sensory motor* stage of development. Most of the learning in this period is driven by biological maturation of neural and body areas, and involves gaining understanding of what sensory input means in relation to objects and other people in the outside world, as well as gaining control of coordinated motor behavior. During this period there is extensive development of fine motor control in the cerebellum. There is also extensive development of specialized cognitive processing areas like Broca's and Wernicke's areas for speech and language processing.

While learning and development in infancy is a fascinating field of study, we will not consider it in any depth. The processes that underlie learning in the ULM are operative at this age, but the child does not yet possess language or other symbol systems that are the currency of school learning. While the ULM accounts for all learning including early years, our intent in this book is to focus on school learning.

The Operation of the Neuron

For all of the biochemical complexity underlying how a neuron works, its operation can be described in simple terms: A neuron "fires" or produces an electrical output in response to having been "fired upon" by other neurons. All neurons have an input end and an output end. The input end can be connected to (receive input from) one or many other neurons. Once this input passes a threshold, the neuron sends an electrical potential that produces release of biochemicals (neurotransmitters) at the output end that are input to the neurons with which it is connected. These connections are called *synapses*.

It is this basic operation of the neuron that defines what it means to "learn." *Learning occurs when the firing ability of a neuron is changed*. This can occur within the neuron by changing the firing threshold or by increasing the amount of input being received. Again, the actual underlying biochemical processes involved in neural learning are very complex, but these neural processes are driven by very simple mechanisms.

Stated simply, neurons are changed by activity. The internal mechanisms of the neuron are changed each time the neuron fires. The more the neuron fires, the easier it is to fire again. Similarly, the connections between neurons are changed with each firing. When neurons that are connected together fire together (simultaneously or in sequence), the connections are strengthened.

At birth, the brain contains massive numbers of neurons and neural connections. As noted previously, the majority of these are specialized neurons that record sensory input and direct motor movement. Regardless of what role an individual neuron plays, however, all follow the same learning mechanism. If it is fired, its ability to fire again is increased; if it is not fired, its ability to fire again is decreased. As the brain matures or "develops", the pattern of neurons and how they are connected is determined by what neurons and neuronal connections fire or don't fire. Neurons grow or die and neuronal connections are created or eliminated based on which are active.

One of the best examples of this is spoken language. Among all languages combined, there are approximately 800 phonemes (meaningful vocal sounds) used. At birth, humans can vocalize all of these phonemes. Any specific language (like English or French) only uses a subset of these phonemes (in English it is approximately 44 depending on dialect). As a child learns her native language, the neural connections that produce the phonemes in that language are strengthened and those not in the language are weakened. Over time, the ability to produce phonemes that are not part of the native language is diminished or even lost. This in part accounts for the difficulty that persons learning a second language in adolescence or adulthood can have with accurate pronunciations and why their speech often has a pronounced accent. Although phoneme neurons can regrow or restrengthen as the second language is learned, they may never fully recover to the level of a native speaker, which is why an adult language learner may never be able to speak without an accent no matter how strong their language fluency.

The Architecture of the Brain

The human brain has an identifiable anatomy. Areas of the brain are in the same location in all people. Most of these areas contain neural groups that perform certain functions. For example, the occipital lobe toward the back of the skull processes visual sensory input from the eyes. Although the gross anatomy of the brain is a result of genetics, the actual neural connections themselves mostly result from learning – that is neurons firing or not firing. So while it may be accurate to say that vision or language are processed in specific regions of the brain, how a region processes information is determined by how the neurons in that region have been strengthened or weakened through past firing. This notion explains why brain surgeons often map the functional brain regions during surgery and especially prior to removing tissue[4]; we're all the same, but we're all different, too. Our brains show a remarkable plasticity. In the event that a procedure ends up with tissue being lost, it remains possible for other tissue to acquire the lost knowledge through subsequent patterns of neural firing using an area of the brain that is not generally associated with that function.[5]

In response to the "nature vs. nurture" argument of whether the brain (or behavior) is primarily due to genetics or biology or due to experience, *we can safely say that it is both*. While the brain does come prewired to receive input from the world in certain ways (the cones and rods in the eyes respond only to certain parts of the light that is in the world and the occipital lobe processes these in specific ways) or produce a defined set of motor movements (the human arm and leg can only move in certain ways but not in others in response to neural signals), the actual things that we see (what we come to recognize as objects or meaningful visual entities) or movements we can do (such as dance or catch a ball) are the result of our experience which has produced specific patterns of neural firing.[6]

It is proper to say that the macro architecture of the brain is genetic, but the micro-architecture is environmental. Because the macro architecture is common to all people, it also is proper to say that differences between people are due to differences in micro-architecture. One person doesn't differ from another because she somehow has a unique anatomical structure that someone else doesn't have. For the most part, one person differs from another because she has a different pattern of neurons and neural connections within each of her brain areas. The vast majority of these differences in patterns are due to learning.

What Is Knowledge?

Knowledge, as we use the term in the ULM, is entirely the result of the micro-architecture of the brain. While it may be possible to locate the processing of something like numbers and mathematics to a rather specific anatomical region of the brain, the ability that one has to do computation, algebra, calculus or other mathematics is not a result of having or not having this anatomical region; everyone has this region. It is due to neural patterns in that region having been strengthened and weakened in ways that correspond to learning algebra, calculus, etc. The strengthening and weakening of neurons is learning. *Thus, the micro-architecture of the brain and as a result, virtually all of our knowledge is the result of learning*.

It also happens that brain tissues can acquire new micro-architectures as the need arises. For example, amputation of an extremity can lead to "reassignment" of the brain tissue that once took care of input/output from the amputated extremity to another task.[7]

How Learning Works

We have stated that learning is the change in a neuron that strengthens or weakens its ability to fire. But, how does this neuronal change happen? We have broadly noted that it happens because the neuron fires, but what exactly determines when a neuron fires?

Each of our sensory and internal homeostatic regulatory systems has a dedicated input area (the occipital lobe for vision for example). These input areas have some dedicated neural storage areas and the neurons in these are fired, and hence strengthened or not strengthened, based on the direct pattern of input from the sensory organ to which they are connected. These mechanisms account for gross sensory recognition and discrimination processes and require no further neural systems. While these neurons are following the same neural learning mechanisms, we are not really concerned with this basic level of processing in the ULM.

We are concerned with the next levels. We receive a tremendous amount of sensory (and internal proprioceptive) input at any one moment. The knowledge present in a sensory input area is extremely low level. It is something like boundary discriminations between light and dark areas that define the presence of an object in the visual field. There is no knowledge of what this object might be in the sensory area. Sensory areas simply aggregate their specific component pieces into an output that is sent out of the sensory area to the rest of the brain.

The place where this sensory output is sent is working memory. Working memory does not have a clearly defined anatomical area. It appears to be a collection of brain regions in the prefrontal cortex along with other structures such as the hippocampus. Each of the sensory (and proprioceptive) input areas generates neural output that is sent as input to working memory. There is too much of this sensory input to deal with at the same time, so a primary function of working memory is to choose what of this input will be ignored and what will be processed. The general term for this is *attention*. *Working memory is where, at a given moment in time, we attend to some inputs and not others*.

Sensory input that is attended to in working memory is a candidate for being stored in the long-term permanent memory neurons of the brain. Most of these areas are in the cortex and are collectively referred to as long-term memory or just memory. Although the full transfer process in the brain is not well understood, input that is attended to activates neurons in a temporary memory area, possibly the hippocampus as well as some others, that create a neural representation of the sensory input in working memory. This information is maintained over a few hours interval through what is known as *long-term potentiation*, a process that keeps the temporary neural pattern active so that it doesn't disappear. In other words, long-term potentiation is a state of storage in between permanent or long-term storage and fleeting or momentary storage. If the neural pattern does not decay, it activates a neural pattern in the cortical region that produces a permanent memory trace of the original input. *As this trace learning occurs in the cortical neurons, their patterns of connectivity become changed*.[8] Psychologists use the term *storage* to describe the process of turning a specific input into a permanent trace.

Meaningful Learning

It is reasonable to ask at this point, "so what?" Yes, a "trace" of the sensory input has been laid down in the cortex, but what good is this? An isolated memory of a single sensory image, say a round object, isn't really very useful. How is this "knowledge?" We will grant that if all working memory did was to create random isolated, stored sensory inputs (images, sounds, smells, etc.), it wouldn't be worth much. But, working memory does much more.

Suppose that the sensory input attended to in working memory was something you had attended to before. In this case, there would already be a neural pattern in the cortex that was the same as this input. When this happens, the neural pattern in the cortex is fired or activated. Rather than create a new memory, working memory "recognizes" the pattern as a known and the existing memory is strengthened. If a sensory input occurs frequently, the strength of the cortical neural pattern can become large and the speed at which working memory can "recognize" it (or match it to an existing pattern) gets faster. The general term for this process of matching working memory input to long-term memory neural patterns and activating them is *retrieval*. American adults usually can decide quickly that a round object is an apple as opposed to another similarly sized object like a baseball, or distinguish a Granny Smith apple from a green tomato.

Now, let's assume that rather than a single sensory input, there were two sensory inputs in working memory. When this neural pattern is stored, both sensory inputs will be part of the pattern. Now at a later time, one of these sensory inputs is attended to. This will trigger retrieval of the pattern, but it won't be just this one sensory input pattern that is activated and retrieved, it will be both of the originals. This is because both were stored together originally. Remember, neurons are chained together so that the activation of one "fires on" any other neuron it is connected to. This chained firing of the one sensory pattern match will activate the second. When you see an apple, you almost certainly also are firing up what you have stored under "sphere."

Let's next assume that one of the original sensory patterns was attended to in working memory, but there was a third new sensory pattern in working memory with it. The original sensory pattern would "retrieve" the two patterns in long-term memory that are connected and because these are now active, the new sensory pattern in working memory would be connected to them. The resulting neural pattern in long-term memory would now have all three sensory patterns and could be retrieved by the presence of any of the three.

This process can chain together sensory input and existing cortical neural patterns virtually infinitely. This is how our knowledge of things and concepts are built. The entire process runs by *pattern recognition*. We "understand" what something is because a pattern of sensory input in working memory matches a pattern of neurons in cortical long-term memory. When one part of a chain is matched, the entire chain is activated because patterns are linked by chaining of neurons. The process of match and retrieval is called *spreading activation*.[9] You see the Granny Smith apple, but you also seem to be able to link the visual input to input from smelling and tasting it.

While we often speak of bits of knowledge being "stored together," as in the case of the taste, odor, and appearance of a Granny Smith apple, it is not necessary that this storage takes place in side-by-side brain structures. They only need to be linked. It has been pointed out that "Knowledge of how distinct brain regions contribute differentially to aspects of comprehension and memory has implications for understanding how these processes break down in conditions of brain injury or disease."[10] There are cases when, as the result of an accident or disease, parts of prior knowledge are lost because the links connecting the knowledge are disrupted.

We have sketched out the building of neural connections and patterns for sensory input. The sight of the apple becomes connected to its taste and smell. The same processes are at work in motor behavior and in connecting motor behaviors to sensory patterns. For example, random movement of an arm becomes a purposeful skill like hitting a tennis ball with a racquet. The arm movements become chained to memory patterns of tennis balls, the court, and the opponent such that the incoming sensory image of the ball in flight retrieves the changed patterns of visual recognition and motor movements needed to hit the ball. We make sounds and learn to sequence them to make words and sentences and ultimately connect them to the complex knowledge of reading and writing. We will consider such learning processes in greater depth in subsequent chapters.

The Centrality of Working Memory

We hope it is becoming apparent why working memory is the centerpiece of the ULM. *Learning doesn't happen without working memory*. If working memory were unlimited and could attend to lots of sensory inputs at once, learning would be fairly fast and easy. We have already noted that there is too much sensory input coming in at one time to attend to all of it. We can't attend to it all because working memory has limited capacity. It can queue up very few things at once: only about four. Also, it can attend to only one thing at a time. *Working memory is the bottleneck of learning*. This places significant constraints on just how quickly large bodies of chained together neural patterns can be created. Much of the ULM is devoted to understanding working memory capacity limits and their learning and instructional implications.

Motivation

The final consideration for learning is motivation. The brain areas associated with working memory have substantial connections with the brain areas associated with emotions. This means that working memory is receiving significant input from emotions. These emotional inputs can affect attention and how the capacity that working memory has is allocated.

Attention requires effort. Try concentrating on something and only that for 30 s. This will become increasingly difficult as your attention will start to drift and you will have to "work" to maintain your concentration. Although attention does not need to be directed and controlled in this way all the time, it is clear that directed sustained attention requires effort. We need students to pay attention in classrooms. *Motivation is the psychological construct that is used to describe those things that* *impel and sustain us in putting forth effort.* In the ULM, we focus on motivation in the context of working memory. Motivation in working memory is derived from emotional inputs as well as from knowledge that has been stored about previous performance, goals, rewards, and ourself. These motivational influences determine the effort level that is put into learning.

General Rules of Learning

This brief overview of the neurobiology of learning can be summarized using five general rules. These form the foundation for the ULM.

1. New Learning Requires Attention

No sensory input is stored in long-term memory without passing through working memory. It is fundamentally true to say that what is stored in long-term memory is what working memory was attending to. *Teaching and instruction, therefore, are about getting students to "attend" to things.*

2. Learning Requires Repetition

Learning is more than just storage. Remember that neurons are strengthened by activation. While it is technically true to say that a sensory input is learned at the moment it is stored in long-term memory, if that long-term memory neural pattern is never reactivated, it will remain very weak and may ultimately decay away to nothing. To create a long-term memory that becomes a lasting piece of knowledge, a specific long-term memory pattern must be retrieved. And it must be retrieved over and over. Knowledge that isn't used is lost. It will not be enough to simply have students memorize (store) a fact. That fact must be used. For example, research has repeatedly shown that when learning new vocabulary, repetition and use in a variety of contexts greatly improves learning.[11] *The educational implication is that teaching needs to include retrieving and for skills, practice.*

3. Learning Is About Connections

What is stored together stays together; what is retrieved together stays together. Knowledge is not contained in a single neuron. Knowledge is a connected pattern of neurons. Higher, more complicated forms of learning, like concept formation or skill development, are essentially about connections. They are about insuring that the things that should be connected are connected and the things that shouldn't be connected aren't. *This means that effective teaching and instruction are about insuring that students are attending to the proper connections.*

4. *Some Learning Is Effortless; Some Requires Effort*

This seeming paradox is actually a byproduct of basic neural learning mechanisms. Sensory and proprioceptive inputs are continuously coming into working memory. Working memory continuously attends to some of this input, stores some of it, and retrieves things based on pattern matches to this input. This ongoing input, attention, storage, and retrieval cycle doesn't require a person's effort. It typifies most "everyday" learning, especially memory for our own day-to-day life. You probably didn't have to exert any conscious attention or effort to "learning" (storing in memory) what you ate for breakfast.

On the other hand, deliberate learning of some specific information or skilled movement is difficult.[12] For example, you may have to reread the instructions to operate the espresso machine you used yesterday for the first time. In these cases, we don't want working memory to attend to just any old thing or what is generally happening to us; we want working memory to attend to a particular thing. This, as we noted previously about concentrating on something for 30 s, requires effort. *Since school is about deliberately learning specific information and skills, learning in school will be difficult*. Transferring what you do learn in school to new applications requires even more effort. Although we want students to like school and be interested in what we are teaching, we should not forget that school is difficult. If learning was effortless, we wouldn't need teachers or schools.

5. *Learning Is Learning*

All neurons learn the same way. This does not mean that all students learn everything equally. There may not be different ways of learning, but students are certainly motivated by different things and aided by different prior knowledge patterns. Because differential motivation will influence working memory attention and produce differential learning, teachers and instructional designers need to pay attention to different ways of motivating and engaging different students. Also, instruction and teaching methods may affect learning differently for students. Learning is first and foremost a product of how students are attending (allocating) working memory. No teaching or instructional method produces exactly the same working memory allocation or connection to prior knowledge for all students. So no single instructional or teaching method can guarantee the same learning results for all students.

You might immediately think to yourself, say, that people from different cultures learn differently. We are asserting that, *at the level of the neuron, human learning is human learning*. We grant all of the things that come from prior experience. It is certain that what we already know impacts what we can learn next or most easily. We do not grant that the cellular details are different in any way. Most important, regardless of our prior experience, new learning requires allocation of working memory.

Our Last Words on the Neurobiology of Learning

A chef creating a gourmet meal doesn't need to deal with the molecular structures of the foods being cooked. Similarly, educators do not need to deal with the underlying biochemistry of neurons to develop effective teaching and instruction. In classrooms, we don't directly contact a student's neurons. (At least we don't yet.) We deal with knowledge and learning in the context of things in the world, both physical and social; language, mathematics, science, physical education, history, and so on. The rest of this book will deal with learning in the context of these things.

But, the way working memory attention and capacity influence how children learn to read or acquire skill in math problem solving, the way knowledge of history or science is understood, and the way students are motivated all are due to the way neurons in the brain work. *Teaching and instruction work when they create conditions that ultimately trigger these neural learning mechanisms in students*. The five rules of learning derived from the micro context of the neurobiology of learning apply also to the macro context of learning in the real world of classrooms.

Notes

1. Miesenbock, G. (2008, October). Lighting up the brain. Scientific American, 298, 52–59.
2. Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). Principles of neural science (4th Ed.). New York: McGraw–Hill. For a more readable summary see: Kandel, E. R. (2006). In search of memory: The emergence of a new science of mind. New York: Norton.
3. Proprioception, a sense of movement and "how things are working together," was first described by Giulio Cesare della Scala in 1557 AD. http://en.wikipedia.org/wiki/Proprioception (Accessed March 22, 2009).
4. http://www.mayoclinic.org/awake-brain-surgery/brain-mapping.html (Accessed March 22, 2009).
5. Doidge, N. (2007). *The brain that changes itself: Stories of personal triumph from the frontiers of brain science*. New York: Penguin (Viking Adult).
6. A tale of a person whose blindness was partially reversed and the disturbing consequences of the new found sensory input – as learning from his past acquires entirely new perspectives. Sacks, O. (1995). To see and not to see. *An anthropologist on Mars: Seven paradoxical tales*. New York: Knopf Publishing Group (Vintage).
7. Florence, S., Jain, N., & Kaas, J. (1997). Plasticity of somatosensory cortex in primates. *Seminars in Neuroscience*, *9*(1–2), 3–12.
8. Squire, L. R., & Kandel, E. R. (1999). *Memory: From mind and molecules*. New York: Scientific American Library/Scientific American Books.
9. http://en.wikipedia.org/wiki/Spreading_activation; http://en.wikipedia.org/wiki/Spreading_activation http://en.wiki/Connectionism (Accessed March 22, 2009); Anderson, J. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 261–295.
10. Maguire, E., Frith, C., & Morris, R. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain, 122*(10), 1839.
11. Atkinson, R. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology, 96*(1), 124–129.

12. Research continues on the neural basis of this distinction. It is clear that effortless memory involves the hippocampus which “may store a memory index” for events. Over time and as the result of effort, different brain regions may overtake the need for this index with something that is more efficient. For a discussion, see Takehara-Nishiuchi, K., & McNaughton, B. L. (2008). Spontaneous changes of neocortical code for associative memory during consolidation. *Science, 322*(5903), 960–963.

Chapter 3
Working Memory

Working memory is central to all learning. This is embodied in the first principle of the ULM: Learning is a product of working memory allocation. As we sketched out in our discussion of the underlying neurobiology of learning, working memory allocation involves attention and capacity limitations. In this chapter, we elaborate how these operate.

Working Memory Capacity

Interest in working memory can be traced to George A. Miller, a psychologist whose distinguished career was recognized by the awarding of the National Medal of Science in 1991. In 1956, Miller published one of the most famous and widely-cited papers of all time: "The magical number seven, plus or minus two: Some limits on our capacity for processing information." In this paper, Miller summarized many studies of sensation and memory and concluded that people seem to have a limited capacity for short-term memory of sensory input, a capacity of about 7 ± 2 things.[1] As the study of memory and cognition began to flourish in the 1970s, numerous scientists started examining Miller's limit. Alan Baddeley and his colleagues were among the most prominent of these researchers. Baddeley and Hitch coined the term *working memory* for this limited short-term capacity.[2]

Current models of working memory generally agree that working memory has two main components:

- A working memory storage area for temporarily and briefly holding elements of sensory input and/or knowledge retrieved from long-term memory.
- A processing system encompassing attention and other cognitive actions that operate on (i.e., change) the content of the temporarily-stored elements.

The working memory storage area corresponds most closely to Miller's original memory capacity. Further study of this capacity has reduced Miller's original 7 ± 2 down to approximately 4.[3] Working memory storage capacity seems to be a function of span and possibly speed:

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_3,

- span, the absolute number of individual elements (things like sensory input or retrieved knowledge) a person can retain at one time, and
- speed, how fast a person can do some type of cognitive action or processing on the elements being temporarily retained before they decay and are replaced with new elements.

Working memory storage can be limited by not having any span left because all available holding places (often called slots) are occupied or by not being able to process everything that is in span before it is erased. Just think about the last time you sat in a professional development session, and the speaker rattled off a series of effective ways to teach students how to comprehend what they read. Unfortunately, she did this without including the slide with those recommendations listed in her *PowerPoint* presentation. You wanted to remember each of them, but there were six intriguing suggestions in her list – an overload for your working memory span. You try to write down what you do remember, but find that you only partially recall the last one – your retrieval for all the words you want to write in your notes is not fast enough to hold them in working memory. You resort to glancing at your neighbor's notes to try to refresh your memory about what you wanted to write.

Studies of working memory processing are recent. As a result, today more is known about how the storage component functions than how the processing component functions. The primary component of working memory processing is *attention* (sometimes called *controlled attention*). Attention generally refers to how many of the elements in working memory storage can be focused on at any time despite distractions and interference. Think of this as reflecting how well you can maintain concentration. Processing largely involves combining temporarily stored elements. Learning is about connections; processing capacity determines how many simultaneous connections can be made between temporarily stored elements. These connections transform or change the original elements into something new. The four elements limit appears to be the capacity for processing or combining in attention or, the same as the number of storage elements.[4]

How Working Memory Functions

At the neural level, working memory is receiving a continuous stream of inputs from the senses. Since working memory can hold only about four elements in its temporary storage, all of the incoming sensory input can't be saved. There are pre-working memory selection mechanisms in the sensory processing areas that influence what does or does not get into the available four slots in working memory. The neurobiology and cognitive operation of these are beyond our scope here. Suffice it to say that they tend to work based on novelty (i.e., patterns of sensation that are new) and salience (i.e., how strong or large is the sensation-like a loud noise). Once something is in working memory's temporary storage, however, the processing component of working memory determines what happens to it.

Think about something as simple and common as being awake in the world. Sensory sensations are continuously coming at you. As soon as one sensory input is put into temporary storage, another different input is waiting to be put there. To accommodate the continuous nature of the real world, working memory only temporarily holds information for a very brief period of time. After this time period, sometimes called a cycle, the temporary contents are "erased" and new input comes in. Without something intervening, inputs would flow continuously through working memory as they were happening. To differentiate or focus on some piece of this input, something in working memory has to stop the continuous flow. This something is *attention*.[5]

When we *attend* to something in temporary storage, the attended element is not erased; it is held in storage beyond the normal temporary memory cycle. New input is blocked from replacing it in temporary storage. If one or two elements are receiving attention, there still is reduced room for new input. If attention were focused on the entirety of temporary storage, however, all input would be blocked. Blocking all new input often proves dangerous; a tree might pick that very moment you are attending to the temporarily stored elements to fall on you. That's why driving a car and speaking on a cell phone is dangerous. Most driving situations require minimal resource allocation. If the cell phone conversation requires many resources, however, those needed for driving may become unavailable leading to an accident. Attention can be redirected from current temporary contents to new sensory input. On the other hand, most of us have had experiences of reading a "page turner," a book we just can't seem to put down, and discover that we have read way past our normal bed times – something for which our bodies pay the price of being very tired the next day.

Recent brain research has identified a buffer area that is a part of working memory that allows the current temporary memory contents to be held while new input enters temporary memory for immediate attention. Working memory can "task switch" between temporary memory and the temporary memory buffer.[6] Think back to when you were reading a page turner – and were interrupted by a spouse or child or roommate. A momentary interruption, perhaps where you exchange a brief greeting, might not disturb you. Responding to a question, however, would most likely cause you to "lose your place" in the reading.

Attention does not just pick from random sensory inputs; it directs sensory input. When sensory information is attended, working memory can signal the senses to focus on the area where that attended sensation came from. If something "catches your eye," working memory will direct your eyes to focus on that and collect more input. Attention to something in temporary storage prioritizes any related sensory information. We experience this focused chaining of sensory input, attention, redirection of the sense, new sensory input, continued attention, new redirection of the sense, and so on, as *concentration*. As this concentration is going on, the blocking of other sensory inputs can be quite profound. Think about reading the page turner. One of the authors can remember being so absorbed and focused on reading a book, that his grandmother kept trying to call him to dinner with increasing volumes of shouting until finally a very loud scream broke his concentration and caused him to be in trouble because he didn't come when called.

Working memory does not deal with just sensory input. Working memory interacts with long-term memory. A key aspect of working memory is that it activates or retrieves knowledge from long-term memory. While taking in and attending to sensory input is a critical function of working memory, activation of long-term memory knowledge is even more important. Our ability to utilize our experience, or even more basically to make sense of incoming sensory input, is dependent on long-term memory activation. If this were not so, then each input would be treated as something new and unfamiliar, and we would never accomplish anything.

Pattern matching is the basic mechanism of working memory and long-term memory interaction. When a sensory input in temporary memory is the same as something in long-term memory, the long-term memory is retrieved. Not only is the specifically matched long-term memory retrieved, but anything connected to it also is retrieved. In this way, working memory connects the outside world to our internal store of knowledge. Sensory input triggers retrieval of long-term memory knowledge, which can then be attended to and processed on by the processing components of working memory. This is how we can see $7 + 9 = \_\_\_$ and retrieve the knowledge from long-term memory that this is a mathematics equation and the answer is 16.

Pattern matching can evoke complex patterns. For example, when a teacher sees a student choosing a book, the student activates one set of knowledge, and the book activates a different set. The teacher is then likely to pattern match the book and the student, asking herself the question, "Is this book appropriate for this student?"

Learning Principle 1: Working Memory Allocation

The ULM holds that learning is a product of working memory allocation. In the ULM, we combine the temporary storage and processing components of working memory into the overarching framework of allocation. For working memory to be allocated to a task, two criteria must be met. Slots must be available for sensory input or retrieved memory. Then, attention or processing has to be directed toward the slotted element. We are using the term allocation in the everyday (rather than strictly scientific) notion of attention as expressed by William James back in 1890:

> "Every one knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called *distraction*, and *Zerstreutheit* in German."[7]

James' description embodies both aspects of working memory capacity, namely selecting elements that are temporarily stored and directing attention and processing to the selected elements.

In the school classroom or other educational setting, it is impossible to clearly separate the temporary storage and processing components of working memory. When we tell a student to "pay attention," we want them to both take the instructional information into temporary storage and attend to it or process it in ways that get the processed information stored in long-term memory. When we say in the ULM that the key to learning is working memory allocation, we are expressing the need for *both* storage and processing of that material in working memory. The way working memory allocation produces learning is embodied in three of the five rules of learning set forth in Chapter 2.

Rule 1: New Learning Requires Attention

The first way working memory allocation produces learning is through attention. It appears that when working memory processing attends to something it has temporarily stored, it is more likely to be permanently stored. That is, it is placed into long-term potentiation and then perhaps into long-term memory. Table 3.1 comes from think-aloud protocols of individuals reading. In this study, the participants were asked to recall information from an article approximately 15–20 min after the reading had been completed. All are verbatim.

In a think-aloud, the research participant is asked to verbalize anything that comes to mind. Think-aloud researchers have argued that think-aloud verbalizations reflect what is being attended to in the current contents of temporary working memory. What is notable in these short segments is how close the recalled answers were to what was verbalized in the think-aloud during the actual reading. This seems to indicate that attention, by itself, is sufficient for later recall. Very recent research has shown that neurons activated during learning are reactivated during recall, strengthening the view that what was stored by attention is what is retrieved.[8] We have no

Table 3.1 Think-aloud and subsequent recall

Think aloud statement	Recall
The father seems to ask more for someone to do the task and the mother seems to want . . . she'll just do it herself	The mothers were more apt to doing(sic) the tasks themselves and the father(sic) were more willing to ask the children.
Mothers take on more housework than fathers.	Mothers do a lot of the household tasks.
Hum, what kind of perceptions do kids have of household jobs? How come just 3, 6, and 9?	This article deals with children's perceptions of household tasks during grades 3, 6, and 9.
So regular jobs, sometimes jobs, and never jobs.	3 categories: tasks they often did, tasks they sometimes did, and tasks they never did.

way of knowing how long this memory was retained. The 15–20 min interval was within the time frame of long-term potentiation, so it is possible that no long-term memory was created.

Rule 2: Learning Requires Repetition

Working memory allocation can also increase learning by repetition. In the preceding think-aloud example, attention was focused on some information once. What if attention were focused more than once? Attention keeps a temporary memory element from being erased. A single act of attention, however, can only hold an element for a few temporary memory cycles. What if we kept repeating our attention focus? This would keep putting the element back into temporary storage over many more cycles. The term for this in the learning and study strategies literature is *rehearsal*. In a think-aloud, we would detect this as someone repeating something to themselves over and over. We don't quite know exactly how rehearsal works at the neural level. Is it just time in temporary storage that matters, is it repetition in temporary storage that matters, or is it actually retrieval and restorage in long-term potentiation that matters? Whatever the details of the mechanism, we know that rehearsal creates permanent long-term memories. This is the foundation of *rote memorization*. Information that is attended to repeatedly is stored in long-term memory. The problem with rote memorization as a learning strategy is that it usually short changes the learner. Those additional connections likely to facilitate retrieval in similar contexts are not made. Transfer to related or novel contexts at a later time is much less likely.

Repetition does not have to be continuous. Although we think of rehearsal as repeating things over and over continuously, the rehearsal effect appears to operate for the duration of long-term potentiation. Nuthall and Alton-Lee used an extensive array of observational and tape-recorded tracking measures to follow children during school lessons, including out-of-school times.[9] They then traced back end-of-the year final exam answers to the lessons where they were being taught. They found that for most information to be answered correctly in the final exam, it had to appear a minimum of four times in verbalizations or other activity that indicated the information was being attended to. They further found that information needed to reoccur at least once within a day to be retained. In the language of the ULM, the information had to be brought into working memory multiple times within contiguous days to be learned. These repetitions likely involve retrieval or reactivation of the long-term potentiation areas, with the multiple repetitions sufficient to transfer the long-term potentiation to long-term memory.

The final way repetition works is over an extended time period. When working memory processing attends to some sensory input and activates/retrieves knowledge from long-term memory through pattern matching, the long-term memory knowledge that is retrieved is strengthened. (A neuron's firing potential is increased every time it is activated). Extended repetition leads to faster and more consistent retrieval of the information. This is common practice; think of learning the multiplication tables, learning lines in a play, or memorizing a chemistry or physics formula. While at some level, these are "rote memorizations," they may be highly meaningful. Extended repetition makes them more accurately recalled and more usable.

Extended repetition is critical for retrieval. Knowledge is retrieved based on pattern matches. To be usable or to transfer to different settings outside the classroom, even rote knowledge, like a math formula or economics fact, must be connected to a variety of retrieval cues. An example of this comes from elaborated repetition based on work done in primary grade classrooms, where word learning is crucial in laying the groundwork (or neural connections) for later complex knowledge development. Beck and McKeown advise using an instructional progression from their *Text Talk* project that begins with contextualizing a word by talking about its use in the story ("In the story you will hear about how the farmer trudged home after a long day of work.[10] That means that he walked heavily, like he was really tired.") Then, the teacher explains the meaning of the word. ("When you trudge, you walk like your legs are tired from working and they feel heavy.") Next, the children repeat the word, articulating it clearly to form a clear mental representation (Say trudge with me: trudge.") Subsequently, to think about transfer, the word is used in contexts that are different from the book ("You may trudge back to your tent, if you have been out hiking all day.") ("You might trudge if you walked through a path with a lot of thick, heavy mud that stuck to your boots.") Teachers would then push up the level of students' thinking, by having them make a few evaluations using the word ("Would you trudge to the kitchen, if you just finished taking a nap and were full of energy? Why or why not?") The students would then offer examples of their own ("I trudged home after I . . .). Finally, the word's mental representation would be further supported by asking a question that incorporates its meaning. ("What word means to walk heavily, as if you were exhausted?") Within this instructional format, students hear and say the word repeatedly in ways that strengthen their long-term memory knowledge of the word and increase the number of visual, auditory, and contextual retrieval cues for pattern matching.

Rule 3: Learning Is About Connections

Working memory doesn't just focus attention; it also processes information. When processing occurs, the result is likely to be stored in long-term memory. Craik and Lockhart were the first to note what they called "levels of processing."[11] Similar concepts emerged in the literature with labels such as deep processing, transfer appropriate processing, encoding specificity, and deep versus surface learning strategies. All of these frameworks were based on a common property of working memory: information in working memory that is changed or altered is better remembered. Contrast this with the repetition effect. In repetition, the processing attends to something that is in temporary memory and stores a copy of it. The temporary memory element is not changed in any way from its original state. When working memory processes temporarily-stored information, however, the original temporary memory element is changed from its original state into something different. Think about the difference between trying to learn something verbatim versus trying to remember the basic idea or gist of it. How often are students told to remember the idea by "putting it into your own words?" Often when students have memorized knowledge with few connections, replacing some words with synonyms operationalizes "in your own words." Any time working memory manipulates temporarily held contents, the result of that manipulation appears to be stored, at least into long-term potentiation. That is, even if it is not permanently stored, it is retrievable for some time (possibly hours) after the manipulation takes place. In the research and practice of study strategies, we see the levels-of-processing concept as strategies directed at paraphrasing and summarizing information that is to be learned. Most recently, in the cognitive load literature, this act of processing has been termed *germane load*. Germane cognitive load is related to how information is presented in instruction. If the information to be learned is presented in a way that promotes appropriate processing of the information, then it is considered to be part of the germane cognitive load for the learner.[12]

Connections are at the heart of Pavio's well-known *dual coding theory*.[13] While auditory and visual information come in through separate sensory systems and are stored separately, subsequent activation of either one leads to activation of the other. This is the rationale supporting some teaching approaches that are used when working with students with specific language disabilities. Students are asked to say the word to be spelled, say each letter name as they write the word, and pronounce the word again as they look at it. The students are getting both visual and auditory input as they learn the word to help facilitate later retrieval when they see it in print or need to write it.[14]

The processing effect ultimately is about connections between things. Information in the sensory input that is coming into temporary storage is connected in a certain way. Working memory processing can break down existing connections as well as create new connections. A paraphrase or summary is a different way of connecting the current information in working memory. Nuthall and Alton-Lee found that remembrance of information was enhanced by it being connected to other information.[15] Most study and note-taking systems, from creating matrices to drawing concept maps, are based on facilitating connections between the information that is being learned.

The capability of working memory processing to create connections also allows working memory allocation to connect new sensory information to existing knowledge. When pattern matching retrieves knowledge from long-term memory, anything that is currently in temporary memory with that retrieved knowledge can be connected to it and stored with it. This ability to connect pieces of knowledge together underlies how working memory allocation can create the integrated knowledge structures talked about in the literature as concepts, propositional networks, schema, production systems, or neural nets. We will have more to say about the nature of these integrated knowledge structures in Chapter 4.

Expanding Working Memory Capacity

We have talked about the temporary storage capacity as being about four elements. We have noted that those elements can be sensory inputs or retrieved long-term memory knowledge. But, we haven't really talked about the exact nature of an element. As a working memory capacity test, try the following:

Read each row. After the row is finished, look away for 5 s. Then, without looking back, recite the row. Check your accuracy.

x g c w
m q p t x r
z p w x m v b t
m t p j w s d l q z

As the number of letters increased, you probably noticed yourself having more difficulty recalling the letters. If you could recall the last row or even the last two rows accurately, consider yourself unusual. Now try this test, doing the same as before. Read the row, look away for 5 s, and then recite the row.

dog farm rocket
onion frame car rodeo

You probably didn't have much trouble remembering the first line and most likely didn't have much trouble with the second line either. But count the number of letters. The first line of dog, farm, rocket has 13 letters. The second line of onion, frame, car, rodeo has 18 letters. In the first test, the maximum number of letters in any line was 10 and you were probably beginning to have trouble with the third line of 8 letters.

From this example, it seems fairly clear that an "element" that is filling a slot in temporary storage is a variable thing. Temporary memory can hold four or a few more letters, but it can also hold four or a few more words, even if these words add up to far more than four letters.

The generic name in the scientific literature for these variable elements is a chunk. A chunk is a connected grouping of knowledge. Accordingly, we can say that working memory has four slots, and that each slot can accommodate one chunk. We will discuss chunks in more detail in the next chapter on knowledge. Suffice it to say for now, a chunk is associated with meaning. Chunks are individual entities (like specific sounds or visual input) that have been connected together into a single meaningful entity, such as letters combining to form a word. The meaningful nature of a chunk can be seen in the following.

Like the previous exercise with letters, look at the following list of numbers, then look away for 5 s, then try to recite the numbers.

6
0
2
5
7
1
2

Fairly difficult. Now look at the following.

602-5712

This seems easier to remember now. What is the difference here? The list of numbers is the same list of numbers, but the second is arranged in the well-known form of a telephone number. Many telephone numbers are meaningful; think of your home or work phone number. Also, we routinely have to deal with phone numbers so we are well practiced in seeing and using numbers in this form. Essentially, we are used to seeing certain number combinations as a "telephone" chunk. We recognize that the Greek letter π (pi) often represents the chunk 3.141.... In language learning, word families or onsets – sounds before the vowel in a word – and rimes – the rest of the word (m-ap, tr-ap, ch-ap, str-ap) are chunking examples that help us recognize words quickly and spell them with less effort once we understand how they work.

Chunks dramatically expand working memory capacity. We really don't know the limits of how big a chunk can be. We also don't fully know whether they are retrieved into temporary memory as a single group or as serially connected or chained spreading activations or some combination of these. Chunks do dramatically increase the amount of knowledge that working memory has available in temporary memory for processing. As we will discuss in the knowledge chapter, chunks underlie most development of skill and expertise.

Chunks reflect the two-way interaction between working memory and long-term memory. It isn't just that working memory retrieves knowledge from long-term memory. The knowledge that is retrieved actually affects the span and efficiency of working memory processing.

Working memory appears to be a biological entity, something that has a physical realization in the brain. Although the basic capacity of working memory is likely a function of genetics and basic brain architecture, long-term knowledge chunks can expand it. Working memory capacity is not fixed. While we each have about four slots, the capacities of these slots grow as a result of chunking. Dehn has addressed many of the issues connected with measuring working memory capacity.[16]

Working Memory as Consciousness

When we go back to William James' quote (p. 22), he singles out consciousness as a property of attention.[17] It is tempting to think of working memory as the seat of our consciousness, our awareness of ourselves, our thinking, and our behavior. Most likely, this notion is not correct. While we might be able to make a case that anything we are conscious of is in working memory, we certainly are not conscious of everything that happens in working memory. Operation of the attention and processing components of working memory can occur without our self-awareness. Similarly, elements can be moved in and out of temporary memory without our awareness. In a sense, to be conscious of the contents of short-term memory or working memory processing, we have to direct working memory's attention on itself.

Similarly, working memory is not necessarily voluntary. As with consciousness, we can exert some deliberate control over what is being input into temporary memory, where attention is directed, and how we are processing what is in temporary memory. But, these working memory operations can operate without our voluntary initiation and sometimes in spite of our intentions. While we might want to pay attention to our teacher in class, our attention may go back to the scene earlier in the day when our dog was hit by a car or to the funny skit that opened *Saturday Night Live*. Sensory input mechanisms are providing continuous feeds and pattern matching in working memory that are operating continuously to retrieve long-term memory knowledge. These things happen whether or not we are trying to control them.

It is important that voluntary control on working memory allocation is possible, however, because that is what allows education to be possible. Because a student can control working memory allocation, we can set up teaching and learning conditions that help students direct their working memory allocation to specific learning materials. We can, in fact, conceptualize teaching and instruction as the management of voluntary working memory allocation.

Basic Rules of Working Memory

We can summarize this chapter on working memory into the following set of rules.

Storage Rules

1. If something in working memory is attended to, store it in long-term memory (the attention effect).
2. If something is in working memory for multiple cycles, store it in long-term memory (the repetition/rehearsal effect).
3. If something in working memory is processed, store it in long-term memory (the levels of processing effect).
4. If things are in working memory together, store them together in long-term memory (the connection or association effect)

Retrieval Rule

1. If something in working memory is the same as something in long-term memory, retrieve the long-term memory contents (the pattern matching effect).

Notes

1. The beginning of formal descriptions of the amount of long-term memory humans could access at one time is attributed to Miller and the number of items was, at that time, thought to be seven. See Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.
2. The original formal use of the term working memory began with Baddeley and Hitch. (Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), The psychology of learning and motivation. (Vol. VIII, pp. 47–90). New York: Academic Press.) Until recently some researchers sought to distinguish this term from another that had been in prior use – short-term memory. Based upon very recent work, the need for that distinction has been re-evaluated (Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*(6), 1038–1063.)
3. Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology. General*, *136*(4), 663–684.
4. Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sci*ence, *11*(6), 236–242.
5. For a recent review and discussion see: Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience, 30*, 57–78.
6. Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*, 594–598.
7. James, W. (1890). The principles of psychology. The quote cited here is found as part of James' book which is reproduced as part of Classics in the history of psychology. http://psychclassics.yorku.ca /James/Principles/prin11.htm. (Accessed October 11, 2008).
8. Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. *Science, 322*(5898), 96–101.
9. Nuthall, G., & Alton-Lee, A. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal, 32*(1), 185.
10. Beck, I., & McKeown, M. (2006). Different ways for different goals, but keep your eye on the higher verbal goals. In R. K. Wagner, A. E. Muse & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 182–204). New York: Guilford Press.
11. Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.
12. Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4.
13. Pavio, A. (1990). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
14. Slingerland, B. H. (1985). *Basics in scope and sequence of a multi-sensory approach to language arts for specific language disability children: A guide for primary teachers.* Cambridge, MA: Educators Publishing Service.

15. Nuthall, G., & Alton-Lee, A. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal, 32*(1), 185.
16. Dehn, M. J. (2008). *Working memory and academic learning: Assessment and intervention*. New York: J. Wiley & Sons. The issues regarding measurement are potentially economically significant. Changes in standard tests coupled with definitions of learning disability have resulted in important changes in special education funding eligibility.
17. Although obviously a very important phenomenon, consciousness is not well understood. We might ask the question, What is different between being conscious or not conscious, as when sleeping? Francis Crick described approaches to consciousness in his book. Crick, F. (1995). *The astonishing hypothesis: The scientific search for the soul (*paper*)*. New York: Scribner. A recent review of studies of persons under anesthesia suggests that consciousness may involve an integration of information, one that is lost during sleep or while under anesthesia. See: Alkire, M. T., Hudetz, A. G., & Tononi, G. (2008). Consciousness and anesthesia. *Science, 322*(5903), 876–880.

Chapter 4
Knowledge

The second core component of the ULM is knowledge. *Knowledge is everything we know or can do that is stored in the neurons in our long-term memory*. Within the educational and cognitive literature, there are a multitude of terms that refer to knowledge. In 1991, Alexander, Schallert, and Hare attempted to identify and organize the knowledge terms being used in literacy research.[1] They identified 26 distinguishable categories of knowledge terms each of which subsumed a number of related specific terms. It is safe to say that since then, the number of knowledge terms in the literature has expanded.

One of the goals of the ULM is to provide a more parsimonious portrayal of knowledge that is more usable by educators. The road to this parsimony is through the ULM's anchoring in neurobiology. This is where we begin.

Long-Term Memory in the Brain

As we use the term in the ULM, knowledge is the micro-architecture of the brain. Knowledge resides in the firing potentials and interconnections of neurons. We have discussed the role of working memory in storage and retrieval of knowledge. What do the effects of these look like in the actual long-term memory areas of the brain?

As it turns out, we do not know nearly as much about the actual neurobiology of knowledge as we do about the operation of working memory or basic sensory and motor processes. There is little hope, therefore, of being able to directly observe the dynamics of the changing neural connections and especially the formation of chunks and broader neural interconnections. The reason that Eric Kandel, who studies these systems, used snails to model neural interconnections was that these animals do learn but have only a few thousand neurons.[2] In contrast, we are thought to have about 100 billion neurons in our brains.[3] We can make certain inferences about what knowledge must look like neurobiologically from our understanding of neuronal learning. The number of neurons in a human brain is vast. We can draw on work in neural network or connectionist modeling to aid our understanding of knowledge. Neural net models are mathematical and computational models (i.e., computer simulations) of neural functioning. They have been developed to help understand the way actual neurons strengthen, weaken, and connect during learning. They also

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_4,

have been used to understand how pattern matching retrieval works. Neural nets also underlie many "expert" computer systems that perform a variety of real-world tasks in science, industry, and medicine.

Neural nets have been especially important in providing a bridge between our macro-understanding of learning and knowledge (derived from a century of educational and psychological research) with studies from neurobiology. Neural nets have shown that virtually all of the findings from educational and psychological research can be realized in a model that functions like neurons in the brain function. The computational and mathematical complexities of neural net models are beyond the scope of this book. The ULM's portrayal of knowledge and how knowledge is learned, however, is based on a combination of neurobiology and neural network findings. Interested readers are directed to Elman et al.'s *Rethinking Innateness* for a reasonably readable description of neural nets.[4]

Storage in Long-Term Memory

Knowledge exists if it has been stored. Storage based on the first rule of learning (attention in working memory) produces a permanent change in neurons in long-term memory. This storage is at the chunk level. Sensory input in working memory cannot be represented with a single neuron. Instead, it will always be some integrated cluster or network of neurons that form a chunk. Once this chunk is stored, the neurons in the cluster that form the chunk will all have similar higher potential for firing. Depending on how this chunk was stored, the chunk may be permanent. Integrated sensory patterns, like specific sights, sounds, smells, and tastes, can be recalled after extensive time periods. A chunk that was stored with an emotional connection (which we will discuss later) can exist permanently even if not subsequently retrieved for many decades. For example, the smell of cotton candy can bring to mind a pleasant trip to the circus during childhood.

Retrieval from Long-Term Memory

As we have already noted, neurons are changed by activation. Activation involves repetition. Neurons in long-term memory change in accordance with the second rule of learning (repetition), just like new information in working memory. In long-term memory, repetition is produced by retrieval. Neurons fire when a pattern match occurs in working memory. The cluster of neurons corresponding to a matched chunk fire; and as a result, this cluster of neurons subsequently has a higher firing potential. Each time the chunk is retrieved, the firing potential increases. The aggregate result is that the more a chunk is retrieved, the easier it becomes to retrieve it. This corresponds to the well-known dictum "use it or lose it."

The repetition/retrieval effect is important because, although storage will produce a chunk in long-term memory, the extent of its enhanced firing potential will remain small if it is not repeatedly retrieved. It may be too small to actually be selectively retrieved by pattern matching. Pattern matching in working memory is rarely exact. It would be rare that a sensory input would be a perfect match to something stored in long-term memory. *What is retrieved is what matches best*. Best is determined by the strengths of the firing potential in the chunks that most closely match. We may see a creature – and retrieve a dog, one we've never seen before, but a composite of many different dogs that we have seen. Non-repeated chunks may technically be in long-term memory, but may never be retrievable. Insufficient strengthening through retrieval may account for much of the "tip of the tongue" phenomenon.

Connection in Long-Term Memory

While we recognize that any meaningful piece of knowledge is a chunk formed by an interconnected cluster of neurons, there is a broader aspect of connection in long-term memory. Think about a ball; any ball will do. You probably get a "picture" in your mind of something like a baseball, football, or soccer ball. Now, think about the ball, and close your eyes and let your mind wander for a few seconds.

You probably started to get a series of thoughts. Maybe you thought about ball games you had played recently or as a child, or games you had watched on TV or at a stadium. Maybe you thought about friends you played or watched with or playing "catch" with your parent. Maybe you thought about playing with your children. Maybe you thought about a large dance event and the gown you wore. Maybe you remembered positive or negative feelings. Our point here is that whatever ball you chose to think about, there were a lot of other thoughts that seemed to be connected to it.

While a chunk is one kind of connected grouping of neurons, there are other types of connections in long-term memory. The basic distinction between a chunk and other connections is that a chunk is a single entity. It might be a relatively small thing, like the visual image of a ball. Or it might be something quite large like our understanding of what a bird is or what "balls" in general are. We often refer to these larger chunks as concepts. But, regardless of how complex these chunks/concepts might be (our concept of bird could be quite complex after taking an ornithology class), they define something that is a unitary whole – a single thing.

One of us uses a task where we ask someone to say quickly as many names of animals as possible. Invariably that retrieval comes in waves. Usually the person says the name of an animal, for example lion, and then names soon thereafter but in rapid succession other animals that are connected to the first in long-term memory such as other cats often found as zoo animals (tiger, leopard, cheetah, etc.). Then there will be a pause as an animal from a different category is activated, and then another burst of names of animals of that type.

Connections can exist between these separate chunks as when you were thinking about your ball earlier. Ball could be connected to another memory of a specific ball game or of friends. Ball might be associated with these, but clearly these are not part of the "ball" chunk. These other memories are not balls.

The micro-architecture of the brain, our knowledge, is made up of these connections. We interconnect specific information about individual things into larger and larger chunks. We interconnect chunks into associated groupings sometimes called *schema* in the literature. *The building of these connections is basically what school is all about*.

The Location of Knowledge in the Brain

We have noted that, in the biological sense, long-term memory results from groups of interconnected neurons that fire concertedly and which themselves are connected to other groups in a broad interconnected association network.[5] Consider the problem of hearing someone speak on the phone. Usually the listener has two problems. One is to decide who is speaking, and the other to comprehend what is being said. Recent studies show that these processes are performed in separate brain regions.[6] When you decide who is speaking, the meaning to you of what is said may change drastically based upon your relationship with the speaker.

Knowing the biological facts about the separated storage of knowledge that we deal with in the context of single, coherent events, however, probably isn't very helpful for writing a lesson plan or determining how to teach algebra to a particular student. Fortunately, we can draw on educational and psychological research that has looked at both functional and structural distinctions between different kinds of knowledge and memory and learning issues surrounding them to guide us.

Episodic Memory

One of the oldest and most basic distinctions in memory research, both cognitive and neurological, is between episodic and semantic knowledge. *Episodic knowledge (episodic memory) refers to our knowledge of our own lives*. It is sometimes called *autobiographical memory*. We have mentioned a key aspect of episodic memory in the fourth rule of learning, namely, learning is sometimes effortless but sometimes requires effort. Episodic knowledge may be learned without effort; it just happens.[7] Keeping an account of what we are experiencing is something that has obvious significance from a survival perspective. If we couldn't remember what we were doing or where we were at when important things like finding food and shelter or avoiding predators took place, we wouldn't be alive long.

Episodic knowledge is attended to automatically in working memory. Episodic knowledge can be very detailed. It contains very specific sensory information such as sights, sounds, and smells. It can be strongly connected to internal sensations such as pain and emotional feelings. It also is highly situated. Our lives happen in specific places at specific times, and episodic knowledge is highly contextualized to those environmental surroundings. While episodic knowledge would form a highly elaborated chunk in the ULM framework, the chunks formed by episodic knowledge do not appear to be long lasting. If I ask you what you had for dinner last night, you will probably have no problem answering. But, how about what you ate two nights ago? A week ago? How about a year ago?

As we move farther back in time, our episodic memories become more and more difficult to retrieve. Unless one year ago there was a special occasion, like a birthday or an anniversary, you are not likely to remember what you had for dinner at all. You will also notice another thing about trying to recall older episodic memories. The farther back you go, the more you actively have to try to reconstruct the event. For example, try to remember what you had for dinner a week or two weeks ago. You probably can't just recall this directly, but you can find it by trying to reconstruct your day. You might recall, "I was at work, then I stopped at the store, then I. . .." With a little work, you might be able to get to the meal. At longer time frames, this reconstruction becomes more extensive. You may not even be able to retrieve specific memories, but instead start to rely on recall of things you typically do, or what the "usual Tuesday" looks like.

Some excellent evidence about the working of episodic memory comes from studies of so-called eyewitness testimony. Among the oft-held pieces of folk wisdom is that eyewitness testimony is of high quality.[8] In fact, clever attorneys often successfully manipulate eyewitness testimony. In an experiment when the word "smashed" was used instead of, say, "contacted," subjects a week later were more likely to report seeing broken glass at an accident scene even though none was shown in the test video.[9]

We can understand why episodic memory works this way if we consider the first two rules of learning. Episodic memories in all of their detail are stored because they are attended to, apparently automatically, in accordance with the first rule of learning. The second rule of learning is that, unless a memory (knowledge) is repeated, its strength will decay. By its very nature, episodic memory doesn't repeat. Remember, it is situated to specific time and place. As one "never steps in the same river twice," one is never again at the same place and time in his or her life. So episodic memories decay because they aren't repeated.

However, your life does have stable elements. Barring some rearranging of furniture by your roommate, your residence will have the same furnishings in the same place this evening as they were this morning. The furniture will have this arrangement every day for the immediate past and future. By the second rule of learning, these repeated commonalities of experience will produce a lasting memory. Essentially we create an episodic chunk that reflects the repeated elements of our experience. These lasting commonalities are what have variously been called schema, scripts, or habits. Our daily lives are ordered primarily because our episodic experience may be represented as a "routine" where events and objects have constancy across time. The permanence of these schemata can be seen by our ability to recall what our residence was like when we were children, even if that was many years ago.

The permanence of these episodic schemata does not mean that we can actually recall specific episodic experiences from the past. What they allow us to do is reconstruct past memories, like recalling what we had for dinner 2 or 3 weeks ago. When we recall our past, we primarily build it out of our general schema. Elizabeth Loftus has examined this reconstruction.[10] It is actually quite easy to reconstruct episodic memories of things that didn't really happen to us.[11] This is because we have lost the original specific detailed episodic memory itself and are constructing from more general schema of our life.

We enjoy sharing stories of our lives with others, and these are based on episodic memory. We can relay a rich picture of significant moments in our lives and allow ourselves to "relive" them. Moreover, we can create mock episodic memory from what has been called "vicarious experience" when we read an engaging, rich novel or watch a movie. As a result, Solomon described literature as a way of educating our emotions.[12]

Semantic Knowledge

Semantic knowledge broadly refers to all of our knowledge (memory) that is not about our own life. In other words, all knowledge that is not episodic. Although this might seem like a rather broad class of knowledge to be labeled semantic, the distinction between semantic and episodic knowledge originally arose in literacy research and the term was drawn from language in that field.

Semantic knowledge is what we teach in school. It is the primary knowledge focus of the ULM. In the school context, semantic knowledge often goes by other names such as discipline knowledge, domain knowledge, content knowledge, subject matter knowledge, or simply the specific subject: reading, writing, literacy, math, social studies, science, and so forth. It is broadly reflected in the categories of Bloom's Hierarchy: factual knowledge, conceptual knowledge, application, analysis, synthesis, and evaluation.[13]

The common underlying characteristic of all semantic knowledge is that it isn't about us. While semantic knowledge may be situated in or about specific contexts, times, and places, unlike the situating of episodic knowledge, this situated context is not our personal experience. Semantic knowledge is not about either us or our lives; $2 + 2 = 4$ for everyone. Because of this, semantic knowledge does not have the automatic attention mechanism that seems to be true of episodic memory. Focusing attention on semantic knowledge generally requires something beyond simply experiencing it; it requires effort.

Semantic knowledge also lacks the initial rich detail of episodic memory. Episodic memory begins with a richly detailed experience; that is, it begins as a very large chunk. Most of the situated richness wanes over time, however. Basically, it is a chunk that gets smaller and smaller and ultimately is merged into an aggregate script. Semantic knowledge tends to begin as single or isolated pieces of information, like a single fact about history, such as Columbus landing in the Americas in 1492. With time, this is built up into larger and larger chunks and integrated networks of chunks, such as a narrative about the settlement of the Americas.

Both the need to focus attention deliberately and the need to build chunks from the bottom up require cognitive effort. The fourth rule of learning can be elaborated as *learning episodic knowledge is effortless but learning semantic knowledge requires effort*. Because learning semantic knowledge requires effort, motivation becomes a key element in semantic learning as we will discuss in detail in the next chapter. Semantic knowledge is initially part of an episodic experience. We will have an episodic memory of sitting in the classroom, reading the book, doing an activity, watching a demonstration, and so forth. This episodic memory of our experience of learning as part of our life is not itself the semantic knowledge that an educational activity is trying to teach. While the semantic knowledge may be part of an episodic memory, if that semantic knowledge is going to be learned, it needs to be distinguished and separated from episodic connections. If it isn't, then it will just become part of our general episodic schema about, say, Mr. Jones' third-grade class. We will remember being in the class but not the actual semantic content that was taught.

Perhaps, more importantly, strong emotions may be connected to episodic memory. If we have negative episodic emotions about school in general, or our experience in certain classes in particular, we may connect these to the semantic knowledge being taught in those classes. When we don't like a subject or have something like "math anxiety," it probably isn't because of anything about the actual semantic knowledge of math (or some other domain). It is more likely because we had a bad or upsetting emotional experience in the class when that subject was being taught. Our bad episodic experience gets connected to the semantic knowledge and keeps being retrieved each time we access this semantic memory. Again, we will discuss this later in Chapter 5.

Declarative and Procedural Knowledge

Two kinds of semantic knowledge have been differentiated in the cognitive literature: *declarative* knowledge and *procedural* knowledge. These are sometimes referred to as "knowledge what" and "knowledge how." Generally, declarative knowledge refers to what we would commonly think of as facts and concepts. Our general knowledge of what things are and what they mean. Procedural knowledge refers to "behavior" both in the sense of muscle movements and in the cognitive sense of thinking. While both declarative and procedural knowledge follow the basic rules of learning, there are meaningful differences in how these operate.

Declarative Knowledge

We can split declarative knowledge into two broad categories: objective knowledge and symbolic knowledge. *Objective knowledge* is knowledge of objects and actions in the outside world as we experience it through the senses. Objective knowledge is the long-term memory representation of sensory input. We can think of the "truth" or accuracy of objective knowledge as being how accurate a copy of the object or action we have stored in our long-term memory. It is easy to think of this as just rote memory. However, remembering sensory experience accurately is not a trivial task. As we noted, many studies of "eye witness" testimony in law have shown that basic rote memory of an event can vary widely across multiple witnesses.[9] As much as we might talk in education about students "constructing" their own meaning, there are many occasions where gaining a very accurate representation of an actual object or system in the world is a legitimate and meaningful learning goal. We would all certainly want our physician's diagnosis to be based on an accurate knowledge of the properties of body organs and observable symptoms, for example.

Declarative knowledge is also our store of socially transmitted *symbolic knowledge*. It includes the meaning of words, moral codes and laws, and our theoretical explanations for phenomena in the world. It includes the laws of physics, as well as knowledge of mathematics, art, music, and literature. These are the things we learn primarily in school or other formal educational settings. Symbolic knowledge technically comes into working memory through the senses, at least initially. Socially transmitted knowledge, as Vygotsky recognized, is not "sensory" knowledge. It is not a copy of a visual, auditory, or other sensory stimulus.[14] The meanings conveyed by the words, pictures, or observations are based on symbol systems like language and mathematics. The "truth" or accuracy of symbolic knowledge derives from rules and relationships within the symbol system rather than from any object or action in the outside world.

Even our objective knowledge is both sensory and symbolic. We simply don't have much, if any, knowledge of the world that isn't paired with a symbolic, usually language, definition. Our sensory knowledge of what a tree looks like as a visual memory (and maybe smells and feels like as other sensory memories) is coupled with the word "tree" which labels the object, along with a host of other verbal descriptors and labels for colors and smells and so on. One of the first learning theories to note this combined sensory and symbolic nature of objective declarative knowledge was Pavio's "dual coding" theory.[15] Pavio showed that knowledge that had both visual and language representation was learned better if the knowledge was presented in both forms to the learner.

Both objective and symbolic knowledge are learned through allocation of working memory by the first three rules of learning: attention, repetition, and connection. The basic unit of declarative knowledge is the chunk.

Building a Chunk

A memory chunk is an interconnected knowledge unit that we imagine as taking one "slot" of available working memory. In one study of memory, researchers took an original visual pattern and made various distorted copies.[16] Participants each were shown several of the copies, but not the original pattern itself. When then given a group of pictures and asked to identify which they had seen, they most often picked the original pattern even though they had never actually seen it. Why would this happen? If we think about how neural learning works, we have an answer. Because the alterations only occur on one or two copies, the repeated common elements in the copies are those from the original pattern. So elements of the original pattern keep getting strengthened while the alterations, which don't repeat, are consequently weakened. The resulting memory chunk for the picture will actually be more like the original pattern than any of its copies.

Think of teaching a child about trees. All trees have some things in common, but each individual tree is unique. As we show a child more and more trees, the common things will become strengthened and the unique things will become weakened. As a result, the child will come to have a chunk for tree that reflects what we might think of as the "essence" of treeness: the underlying distinguishing characteristics of trees. The philosopher Plato wrestled with this question. To Plato and subsequent philosophers, it was unclear how sensory inputs, which were all very unique, could possibly produce a true knowledge of something. Plato proposed that we were born with the knowledge of the "forms" or essence of things. Other philosophers, such as Aristotle, proposed various "inductive" mechanisms that extract this essence from the sensory input.[17]

These ancient philosophers, however, did not know anything about how neurons learn. They did gamble, but they did not understand much about probability – the laws of chance. They certainly didn't understand how neurons implement probability laws. Probability is based on frequency counting. Things that occur with high frequency have higher probability. If we knew the actual frequencies of things, say what percentage of dogs have long hair versus short hair, we could predict exactly how likely they were to occur. We might venture an informed guess about how likely it will be that the next dog we see will have long hair. Since we can't actually see all of the dogs in the world and count the long and short haired ones, we have no way of knowing the exact frequency. Things are never this simple. We might be in a neighborhood where homeowners have a special preference for, say, short-haired dogs.

We could take a local "sample" of dogs, count the long- and short-haired in the sample, and then estimate that the actual frequency would be close to what we observed in our sample. This logic underpins most research in psychology and education; we use the characteristics of a sample to infer the characteristics of everyone. If we took a different local sample, we would get slightly different frequency of long- and short-haired dogs. We could use this to "update" our overall count. A mathematical theorem known as "the law of large numbers" says that, as we take more and more samples, the cumulative frequency distribution of the samples gets closer to the actual frequency distribution.

Where we live in Nebraska, our children have a significant sample of thunderstorm and tornado experiences. Because the frequency of these events is fairly high, they have good understanding of storm phenomena. They have very weak and unique representations for earthquakes, however.

While this discussion of probabilities, frequencies, and samples may seem unrelated to how a chunk gets built, it is relevant because neurons operate according to the law of large numbers. Let's return to trees. Each time we take in a new sensory input of a tree, the tree chunk is retrieved (activated). We can think of this as a sample. The "essence" of trees is the accumulated connected experiences in long-term memory. The frequencies of all characteristics of a tree in the chunk are strengthened (counted) or weakened by whether they are present or not in the new sensory input "sample." As we repeatedly get new sensory inputs of trees and keep retrieving the chunk, we are taking another sample and adding to the counts of some characteristics based on the current tree sensory input. The neuron chunk is keeping the equivalent of a running total of the cumulative frequency of all characteristics. The law of large numbers says that, as we take more and more of these sensory input "samples" of trees, the closer our neuron chunk cumulative frequency gets to the actual frequency of these characteristics in all trees in general. The term general is restricted to one's world of experiences (whether personal or through different media). That is, if you live in the desert and see only short, dry trees, a first encounter with a giant redwood may register very differently. This is why children who have limited experiences have weaker and fewer knowledge chunks and often find reading comprehension difficult.

This applies to symbolic knowledge as well as objective knowledge. Children learn their native language through hearing it used repeatedly. This leads to chunks that contain the basic underlying phonetic, grammatical, and semantic rules of the language. Exposure to the language repeatedly in the environment is sufficient for this learning, even if parents or others do not do some systematic teaching or correcting. Studies have shown that the chunks of language rules children build are actually more rule-based than the language itself. Children will extract the basic rule, such as add "s" for plurals, so well that they have to be taught the exceptions (such as mice for the plural of mouse rather than mouses). The law of large numbers seems to strip away legitimate exceptions to the basic rule as the frequencies are updated in the declarative chunk.

In the ULM, we do not need any kind of special explanation or inductive learning mechanism to explain how a chunk becomes a general concept or rule. We can understand how seeing many different altered copies of a visual pattern becomes a memory of the pattern itself that has never actually been seen. The operation of neurons as a counting mechanism combined with the law of large numbers will produce a general knowledge chunk out of specific sensory input both for objective and for symbolic components of declarative memory. *The way to create an accurate declarative knowledge understanding of anything is to provide enough repetitions of the idea being learned.*

You might well ask, "What if we don't have time to repeat things enough for the law of large numbers to work? The law of large numbers may work great for things we experience every day, like seeing trees or hearing our native language spoken, but in school we have too much to cover to provide extensive repetitions." The answer to this can be gleaned from the notion of a chunk as a frequency count of the characteristics in a declarative concept or rule. The initial exposure to the knowledge, either sensory or symbolic, will establish the initial pattern of strengths when the chunk is stored the first time. The closer this pattern is to the pattern that typifies the knowledge, for example, the closer the altered pattern is to the original pattern, the more accurate the chunk will be. That is, the more like the "original pattern" your declarative knowledge chunk will be, the less alteration through repetition will be needed to get close to the real frequencies. *From an instructional perspective, therefore, the better the quality of the initial teaching or instructional materials (that is, the closer the information in them matches the traits of the concepts, facts, and rules we want students to learn) the faster students will create accurate declarative knowledge chunks*. They will need repetition, but they will need fewer repetitions to create a representative chunk. The ultimate goal of declarative learning is the formation of integrated chunks.

Procedural Knowledge

As described, declarative knowledge can be complex. It allows us to recognize things, make judgments, classify things, discriminate differences, and identify similarities. These types of recognition and classification capabilities can be quite complex. Regardless of this complexity, declarative knowledge doesn't produce an action either in behavior or in thought.

Procedural knowledge is knowledge that produces action. The most obvious and observable action is a motor behavior. Procedural knowledge is what directs motor outputs that produce behavior. Action, however, includes purely cognitive outputs. We experience some of these as thought. Other cognitive actions include retrieval from long-term memory. Declarative knowledge is retrieved by spreading activation from a pattern match to the contents of working memory. But declarative knowledge can also be retrieved by the action of procedural knowledge.

Declarative knowledge simply exists as a memory waiting to be retrieved (activated). Procedural knowledge is purposeful or goal directed. It directs certain cognitive events and motor behaviors to happen, and that accounts for the description of procedural knowledge as "knowledge how." Procedural knowledge is how we are able to do things.

Although procedural knowledge involves motor behavior, it isn't simply muscle movement. We have alluded previously to infant development being a period of gaining motor control; the ability to coordinate our body movements in space. Just as declarative knowledge is not simply a sensory input, because it involves connecting sensory input to symbolic meaning, procedural knowledge is not simply a movement. Procedural knowledge may involve creating knowledge structures that turn movement into coordinated action sequences to achieve specific results. Also, procedural knowledge may involve creating coordinated sequences of cognitive actions or thinking that do not have specific physical movement as their outcome.

A good way to conceptualize procedural knowledge is to think of it as an "If–Then" relationship. If something is present, then do something. If a ball is coming at my head, then raise my hand and catch it (before it hits me). In the cognitive science literature, these are sometimes called "Condition–Action" rules. If some condition is present (it is raining), then take a specific action (open an umbrella). Note that the "If" part of procedural memory is a declarative knowledge chunk. Procedural knowledge is the way that declarative knowledge gets used. When we want a student to learn a skill such as solving an algebra equation, or balancing a chemical reaction, we need to turn her declarative understanding of the rules involved into procedural knowledge.

Examination of procedural memory was actually one of the earliest topics studied in the psychology of learning. In those days, procedural memory was studied as a stimulus–response or S–R connection. This was the foundation of what became known as behavioral psychology. Because behaviorists lacked the tools to examine knowledge at the cognitive and neural levels, they focused on observables, connections between observable sensory stimuli (sights, sounds, and so forth) and observable motor and physical behaviors. Because these observables did not require consideration of "inside the head" thoughts, behaviorists conducted much of their research with animals where experimental research conditions could be better controlled than in research done with humans. Behaviorists performed some of the most famous studies of learning. Pavlov won a Nobel Prize for showing that dogs could learn to salivate to the sound of a bell if the bell was paired repeatedly with food.[18] His learning paradigm is known as classical conditioning. B.F. Skinner studied what was known as operant conditioning and showed that the rate at which rats (or pigeons) would press a bar for food was dependent on whether presses were rewarded (Skinner called it reinforcement) with food every time or randomly.[19] He also showed that animals could learn to press or not press depending on whether certain stimuli (like a green or a red light) were present.

While these studies never dealt with cognitive issues directly, they did illustrate the same principles. Animals "learned" to salivate, press bars, or run mazes as a function of repetition and connection. To get dogs to salivate, Pavlov had to have the bell sound just as food was given, and he had to repeat this connection. The amount of salivation to the bell (strength of the learning) increased as repetitions increased. If the bell was no longer paired with food, the bell would continue to produce salivation for a while, but the amount would decrease over each repetition until it stopped. You may already have noticed that these descriptions from behaviorism seem very similar to our description of neurons learning. Brain researchers have confirmed that neural learning follows the same processes as classical conditioning. The same neural learning processes underlie discrimination learning, where an animal learns to respond when one, but not another, stimulus is present (e.g., press a bar for a red but not a green light). While the ULM accounts for classical and operant conditioning, cognitive research has expanded our understanding of procedural knowledge far beyond the simple behaviors examined by behavioral researchers.

Procedural knowledge is learned through attention and repetition in the same way as declarative knowledge. In humans, we may be concerned with observable stimuli and behavioral responses, as when we want students to be quiet or do something when the teacher gives a signal. In relation to subject matter, we are generally concerned with learning as knowledge in memory. We are concerned with connecting declarative knowledge to the knowledge representation of a behavior. *What is different about establishing procedural knowledge is that the repetition involves both a chunk and an action.*

Building a Procedure: Proceduralization

The end point of declarative knowledge learning is a memory chunk that accurately reflects the frequencies (probability of occurrence) of the specific characteristics of the knowledge. The end point of procedural knowledge is something else. There is more involved than just a cumulative frequency count. We are not simply trying to store an accurate representation of an object or symbolic rule. Procedural knowledge exists to produce an action that achieves some outcome in relation to the real world around us. *Procedural knowledge is goal directed*.

Think about what is needed to open an umbrella when it rains. First, we need the declarative knowledge to know what rain is and what an umbrella is. Then we need the procedural knowledge to know how to operate the umbrella's opening mechanism. If we fully elaborate this action, we might have something like the following:

If something hits my head and it is water, then it is raining.
If it is raining, then I will get wet.
If I don't want to get wet, then I need an umbrella.
If I don't have an umbrella, then find one.
If I have an umbrella, then find the opening mechanism.
If I have found the opening mechanism and I recognize how to operate it, then open it.
If I have found the opening mechanism and I don't recognize it, then figure out how to open and open.

Something as simple as using an umbrella requires an extensive chain of possible conditions and possible actions that will vary depending on the results of previous actions. Learning to put together a procedural chain is not simply establishing an accurate cumulative frequency count of something. It is establishing a chain of actions that lead to a desired outcome or goal.

In educational settings, we are concerned with complex skills such as reading, writing, and mathematics. We are also concerned with developing skills in problem solving in science, history, and other fields. As procedural knowledge, these skills turn out to be very much like the example of opening an umbrella. We want students to develop chains of actions in response to certain states of the world (conditions) that move through alternative outcomes to acceptable results. Consider reading about characters in a story like Cinderella and her evil stepsisters. You can comprehend the story better if you recognize that you need to compare and contrast these characters' characteristics. If you have appropriate procedural knowledge, you may be able to use a strategy like constructing a Venn diagram to show this analysis visually.

The key to understanding proceduralization is understanding the role of outcomes or results. Let's think about learning to catch a ball. If we see a ball coming at us, this sensory input goes into working memory. From working memory, a pattern match retrieves (activates) the relevant declarative knowledge about balls and movement. Assuming we don't just sit there and let the ball hit us, some kind of action will occur. If nothing else, we will probably reflexively duck or put up our hand. Executing these actions requires that motor neurons be activated to send signals to the muscles. This activation retrieves these into working memory. These motor neurons are firing at the same time that the sensory input and activated declarative knowledge is present. Recall the "working memory association rule." Things that are in working memory together are connected and stored together. A connection is made between the declarative knowledge and the action.

The more this connection is repeated, the stronger it will become. How does the connection that succeeds in catching the ball get differentially strengthened? We can make a variety of responses to the ball. Only some of these, however, will result in us catching it. Since we want to catch the ball, we will try to repeat the actions that result in a catch. As we repeat the successful actions more than the unsuccessful actions, the catch action will ultimately become the one that is most strengthened. This doesn't really require that the neurons somehow "know" whether the action worked or not. It simply requires that the action that works be repeated more often while the declarative and sensory knowledge is active in working memory.

This process works the same for cognitive knowledge. Suppose we want a child to learn that $2 + 4 = 6$. When the sensory input of $2 + 4$ appears, declarative knowledge about what the numbers and symbols mean is activated. An answer is given. Any answer other than 6 will not be repeated as often because it doesn't solve the problem; ultimately the answer of 6 gets strengthened.

This process is the same for more complex types of math and science problem solving, just with longer chains. Consider an algebra problem. In this case, the solution may not be a simple answer; it might require the first step to be simplifying the equations by factoring. There might be many steps, but ultimately there are a limited number of sets of steps that lead to a correct answer. Each intermediate step produces an output that becomes new declarative knowledge that forms the condition for the next step. As the entire chain is repeated successively, all of the connections leading to an appropriate answer are strengthened and alternative steps that do not lead to a correct answer become weakened. If there are several sets of steps that can work, the one chosen that is successful is strengthened – while others that might also work are weakened. *Sometimes we end up strengthening a successful approach when other approaches might be more effective*.

This process of creating procedural knowledge is also one by which wrong procedures can be reinforced even if they are wrong. For example, when students learn to read and use books with controlled vocabulary and illustration, they may learn to read words based on the first letter and the picture – a procedure that will not be effective down the road.

The creation of procedural knowledge involves differential strengthening through repetition of the connection between the condition and a suitable action out of all of the possible actions. Our everyday term for this kind of repetition is *practice*. Like declarative knowledge, this selection of the correct condition – action link can be facilitated by teaching and instruction that facilitates doing the correct action from the start. This is why, for example, golf instructors say that you are better off going to an instructor before you try to play than after you have played on your own for a while. The instructor can get you to swing with the right motions from the start rather than having to deal with both strengthening new, appropriate motions while trying to eliminate existing inappropriate ones, which will want to keep happening until the new ones get a higher strength value.

This also explains why it can be better to have students observe a teacher modeling work through an example, and be closely guided during the initial learning rather than have them work things out on their own. *Establishing and strengthening appropriate condition – action chains from the beginning will facilitate speed and accuracy of learning as they practice*.

Automaticity

The ultimate goal of procedural learning is what is known as *automaticity*. The early stages of proceduralization require active attention in working memory. Let's go back to catching a ball. Initially, we will see the ball. The sensory image will pattern match our declarative knowledge chunks about balls and their movement, and about ourselves and our body positions relative to the ball. We will probably be aware of thinking, "there is a ball coming at me." We may even become aware of thinking "catch it before it hits me." This is because working memory is actively involved in processing the sensory input and retrieving the relevant declarative and procedural knowledge. In school, kindergarteners have to think about making the letter M; sometimes they even subvocalize "up, down, up, down."

This active involvement of working memory has been called *controlled* cognitive processing.[20] The problem with this type of controlled processing is that, while we are thinking "catch it before it hits me," the ball is still moving and may in fact hit us before we can react. Controlled processing is slow in relation to the real world. While this isn't necessarily a problem with executing a math formula, it can be a real problem when we need to take action in response to something happening in the world around us – such as avoiding potential injury from a fast-moving ball.

As we have discussed, one outcome of repetition is faster retrieval. As procedural chains are repeated (practiced) two things happen. First, unsuccessful condition–action links are successively weakened, so that the procedural chain becomes more simple and direct. This will increase speed because there are fewer and fewer "mistakes." Second, repetition increases neuronal firing speeds so that condition–action links are executed faster. Although we don't completely understand how this happens, at some point these more streamlined and speeded-up procedures no longer appear to need active working memory; *they no longer need to be controlled*.[21] We can speculate that this may be a function of the procedure becoming so accurate at achieving its intended outcome that there is no longer any need to check whether it did or didn't work. Its probability of success becomes almost 100%. While kindergarteners have to make each letter very carefully and consciously, we adults pay almost no attention to the way we form our letters – we just write.

A procedure that has achieved automaticity can more or less run by itself. Automaticity dramatically increases speed of the action. Essentially, once the initial sensory input in working memory pattern matches and retrieves the first declarative condition chunk, the entire chain in the procedure fires without further working memory involvement. At this point the action becomes what is often called habitual or a habit.

The important point here is that automatization frees up slots of working memory that are not being used for those tasks that once did take up slots. That is, we needn't say to ourselves "up, down, up, down" as we write each M. Yes, you can walk and chew gum at the same time.

When two activities require controlled processing at the same time, the outcomes are different. In spite of the claims often made by student "gamers" and believers in multitasking, we are not really able to divide our attention between two things at the same time. Working memory can switch back and forth; it cannot deal with two controlled processes at once.[22] Therefore, it's a good choice when state legislatures pass laws concerning the use of cell phones while driving. If the cell phone conversation starts to demand working memory allocation, the driving may become unsafe.

When procedures become automatized, however, it is very difficult to apply any new learning or changes to those procedures. Learning requires active attention in working memory. Once started, automatized procedures run independent of attention in working memory. They often require sensory input, but they may not need this to be continuous. An automobile driver can attend to a conversation or listen to the radio. A baseball player chasing a fly ball, however, can't (or at least shouldn't) simultaneously think about world events. If driving conditions deteriorate as the result of conditions like heavy traffic or freezing rain, drivers often turn down their radios and ask passengers to hold their thoughts. Experienced passengers, especially those who also drive, often do this without being asked.

To change an automatized procedure, it must be "unautomatized." It must be brought back into working memory attention. Ericsson calls practice to change procedures *deliberate practice* .[23] One must deliberately focus active attention during practice to keep procedures from simply executing automatically. If they continue to execute automatized, nothing will really change. Instead, the existing procedure will keep becoming strengthened. Tiger Woods, a world-famous and eminently successful golfer, became convinced that the nature of his stoke needed to change if he ever were to improve his accuracy. During the months while he was changing his swing, the quality of his game suffered. In the end, he emerged with what golf professionals regard as a better swing.[24] This effort on Woods' part is a dramatic example of deliberate practice. What Woods does every day is to exemplify classic deliberate practice.

Building Larger Knowledge Networks

Up till now, in our discussion of declarative and procedural knowledge, we have dealt with "single" chunks and procedures. While a single declarative chunk might be quite large and a procedure that accomplishes a single action might involve many chained condition–action steps, these can still be thought of as a unitary whole: a single thing or action. In life, however, we rarely deal with only one thing or action at a time. Our homes are made up of numerous objects that we need to recognize and relate to simultaneously using our declarative knowledge. We rarely do one discrete action and then stop and go on to another discrete action. Instead, we perform sequences of actions, like catching a ball and then throwing it to someone. We do drive while carrying on conversations with our passengers.

Connecting declarative and procedural knowledge into larger networks that link chunks together or link procedures into associations that don't lead to automaticity is a function of the third rule of learning: connections. Because it is difficult to structure experiments or observations of the formation of these larger connections, we know quite a bit less about the details of how they occur neurobiologically than we do about the more direct neural strengthening and close association connections in chunks and procedures. This requires a bit more speculation about the underlying mechanisms.

While frequency is a property of these larger associations, the resulting connections do not appear to be based solely on strengthening through repetition. Larger networks appear to require not just strengthening and weakening through frequency, but also an active inhibition or weakening. Instead of just decaying over time, a neural connection is actively decreased. The brain has mechanisms that actively inhibit neural firing and weaken firing potential. Discussion of the neurobiology of these mechanisms is beyond this book and still not fully understood. We draw upon evidence from cognitive and psychological studies of both memory and schema theory as well as neural net simulations. These are discussed in the next two sections.

Declarative Networks

Declarative knowledge is structured in hierarchies. Hierarchical arrangement is key to neural net models. The physical architecture of the brain also shows hierarchical structuring of neural connections. Why is knowledge ordered hierarchically? The reasons seem to include efficiency and parsimony. Say, we have a terrier named Fido. Fido is a dog that is a specific breed, that is a member of the dog species, that is a mammal, and that is an animal. From a practical standpoint, a hierarchical arrangement of these relationships is efficient. For example, if a dog is a mammal, then we don't need to store core mammal knowledge, such as "vertebra," in the dog chunk. When dog is activated, we can get everything in the mammal chunk along with it because of spreading activation through the hierarchy.

This allows more parsimonious storage and access to information common to many chunks. Many other animals besides dogs are mammals. If we store what is common to all of them in the mammal chunk, we only have to store this knowledge once. All associated animal chunks can access the knowledge in the mammal chunk. Otherwise, we would have to store all of this knowledge in each of the separate animal chunks, which would produce considerable duplication and redundancy of knowledge.

How would neural operation produce a hierarchical structuring of knowledge? Consider a single chunk that a child might have for "animal." There are lots of things that animals have in common. As we have already said, the child's animal chunk will differentially strengthen and weaken elements based on repetition. But, at the same time, each specific animal is also reoccurring, so the child is having repeated contacts with, say, dogs and cats. Both are obviously animals based on common characteristics, but both are clearly different based on other characteristics. We don't want the differences to be eliminated. Within a single chunk, these two competing tendencies cannot be realized. We cannot maximize both the common elements and the distinct elements in a single overall frequency count. To accomplish both strengthening commonality and delineating difference, we need to count both. This is doable in a hierarchical arrangement of multiple chunks. The common elements to both cats and dogs can be counted in the animal chunk and the elements specific to each can be counted in separate cat and dog chunks. Cat and dog can access their commonality from the animal chunk and can maintain their difference in their respective individual cat and dog chunks. The connection of the cat or dog chunk to animal is strengthened each time they are retrieved together. The differentiation of cat from dog can be achieved by specific inhibition connections between them, connections that increase the inhibition each time one occurs without the other, thereby weakening the overall association between them. Synaptic connections can increase or decrease the likelihood of neural firing, which in turn is a function of the type of receptors and neurotransmitter employed at the synapse. Strengthening a connection, then, can lead to either an increase or decrease in the likelihood of firing. In fact, though cases differ widely, it is thought that most synaptic connections are inhibitory.[25]

Procedural Networks

There is less discernable networking of procedural knowledge. If we remember that the end point of procedural knowledge is automaticity, it seems logical that we wouldn't want automatized chains becoming too large. Behavior has to be adaptable to the changing environment in the world around us. This would preclude having too much action chained into tightly connected sequences. Rather than organize into larger, associated groupings like the hierarchical or rule-based structures of declarative knowledge, procedural knowledge appears to link into broader action sequences by leaving a “connector” as the end point of a procedure.

It is usually going to be the case that after executing a procedure, we will need to “evaluate” the result. In relation to overt behaviors, we have just done an action that likely altered the environment. Even if our own action didn’t alter it, the environment will have changed while our procedural knowledge was producing the action. We can’t just keep acting without assessing the new environment. For example, we may have an automated procedure for catching a hit baseball. The obvious next action would be to throw the ball. But, while we are catching the ball, base runners are moving, so we aren’t going to know where to throw until we visually rescan the environment. The last condition–action rule in the catching procedure could be something like “If ball caught, then search where to throw.” This would trigger a scan for visual sensory input but the presence of “throw” would also begin a pattern match to a throwing procedure. The first condition–action rule in “throw” would then be something like “If location to throw to has been identified, then begin throw.” So if the sensory scan hasn’t identified any place to throw yet, the throw procedure would terminate.

This type of linking has been implemented in various procedural knowledge simulation programs. Rather than just stopping, a procedure requests a cognitive action in working memory, like getting new sensory input or retrieving some new knowledge as its terminus. Having a procedure’s last action be something that will produce the condition for the start of another procedure allows not only for linked actions but also for adaptation to changing conditions.

Let’s consider how this could happen. If there is nothing that regularly follows after a procedure executes, it is likely that the terminus would simply be general working memory input of sensory information. If certain actions regularly follow others, the initial tendency would be to include it in the procedure’s chain of condition–action rules. For a procedure to achieve automaticity, however, all condition–action components must execute in sequence 100% of the time. So any action that is only in the chain some of the time will ultimately be removed. Much like the linkage rules for similarity in declarative knowledge, the co-occurrence of the automated part and the following action is high enough to warrant making some links. This linkage can be done by establishing a final action in the first procedure that conditionally activates the second procedure. As with larger declarative knowledge networks, this will involve inhibition as well as strengthening through repetition.

Situated Knowledge and Transfer

The building of complex connections is also at the heart of two of the more prominent issues in the field of learning: transfer and situated cognition. We have discussed the notion of situated cognition previously in considering episodic knowledge and how chunks are built. The formation of chunks that represent conceptual knowledge involves removing specific situated components and creating a general representation through repetition, strengthening common components, and weakening specific situated components. But there is a larger issue of situated knowledge and transfer.

Suppose that we are building a concrete patio behind our house. We are trying to figure out how much concrete we need to order. Do we recognize this as a situation where the geometric formula for volume (height $\times$ width $\times$ depth) applies? In another example, imagine we are trying to adjust a recipe for 4 people to serve 6. Do we recognize that this can be solved using fractional mathematics (multiply the recipe for 4 by 1.5)? These kinds of questions have been asked about what we learn in school. Do the things we learn in school, like mathematics formulas, historical knowledge, and principles of science, ever get used outside of school? Does what we teach "transfer" or generalize to the "real world"?

For a long time, the accepted answer to transfer has been no: things taught in school don't transfer. Reference has been made to things taught in school as being "inert" knowledge, present in long-term memory but irretrievable in real-world settings. This has led to calls to situate learning more in the context of the real world where it will be used. Considerable research has shown that job-related skills are highly situated to the particular context of the job. These job situated skills use strategies that have evolved from experience, often by trial and error, within the specific context and that are independent of rather than drawing on formal knowledge learned in school.[26] Of course, situated job skills don't transfer to another job or setting any better than knowledge learned in school.

In addressing transfer and situated issues, a couple of things need to be kept in mind. First, some school-related skills show wide transfer. Most of us learned to read in school and we can decode and often comprehend any written text in our native language in any place, at any time, and on any media. Similarly, we can write on any kind of media and even use totally different methods (handwriting or word processing) to produce printed text. We can use basic mathematics in a variety of settings, to calculate tips at a restaurant, compare prices at the grocery store, and so forth. If we bothered to learn it in the first place, we can recall information about history and geography that can help us understand current events. *While knowledge learned in school can be inert and inaccessible in the real world, it doesn't have to be.*

Most situated cognition examples like to point to real-world contexts. The only time one of the authors has ever seen trigonometry actually being used was by a friend who is a concrete contractor. People like his friend, the contractor, have developed very accurate situated rules of thumb for determining something like how much concrete is needed. Based upon experience, this friend can probably estimate the volume of concrete needed for a 100 foot curving driveway within about a half cubic yard (a cubic yard being how concrete volume is measured and sold); he certainly fits the situated cognition prototype. Because that half cubic yard may be a large portion of his profit, however, he simply can't afford to "guestimate" by rule of thumb. He uses trigonometry functions to work out exactly how much concrete he needs. The notion that workers in the real world don't use formal knowledge learned in school isn't always true.

All knowledge, when used, is applied in a specific situation. Whether we have a general declarative knowledge chunk of some concept or an automated procedure, these always are retrieved in the context of a specific situated environment. Piaget referred to this as *accommodation*, the necessity that knowledge be able to adapt itself to the specifics of the situation and environment in which it is used. That situation may be "real world," or it may be responding to an item on a Graduate Record Examination (if not a "real" world, nevertheless a potentially important one). Chunk building and proceduralization may create knowledge that reflects common repeated generalities, but that knowledge must always be applied to something specific.

Working memory tries to take a pattern of sensory (or other) input and match its pattern with existing knowledge. Pattern matching leads to retrieval of declarative knowledge chunks and procedural knowledge condition–action chains. From the perspective of the ULM, the key to transfer of knowledge is that some aspect of the environment in which the knowledge will be used must trigger a pattern match. If we want children to be able to apply a math formula to solve a word problem, then elements of the word problem need to trigger a pattern match to retrieve the formula and solution procedure.

The seemingly obvious solution to transfer would be to have students use their knowledge in the situations where it will need to be used in the real world. That way, aspects of these real-world situations can be connected to the chunks or procedural condition elements so that these will pattern match later when the situation occurs again. This is the logic behind apprenticeship approaches where students learn in the actual (or a simulation of the actual) real-world context. Likewise, it underlies internships and practica where students apply classroom learning in authentic real-world environments. It also is the logic behind simulation trainers. The flight simulator used in aviation training is so close to actually flying a plane that one can go straight from the simulator to flying. These approaches to instruction are designed to increase the ability of a situation to lead to a pattern match and subsequent retrieval of appropriate declarative or procedural knowledge.

The limitation to these instructional approaches, however, is that within the limited time available in an educational setting, we can never provide direct or simulated apprenticeship experiences for every possible setting where we might want to transfer knowledge. We need students to develop ways to retrieve relevant knowledge in situations and environments that they have never before encountered.

Problem Solving and Critical Thinking

Developing the capability to retrieve knowledge in situations that have never been encountered before often falls under the heading of problem solving, critical thinking, and sometimes even creativity. From the perspective of the ULM, this type of problem solving ability or critical thinking involves identifying or creating pattern matches. If the existing sensory input has not triggered a pattern match, then one either has to keep looking for new sensory input (more information) or has to transform the sensory input into something that does trigger a pattern match. As we already noted, one of the key processes of working memory is to connect input in working memory storage together in different ways. This process of reconnection may create a new organization of the elements in working memory that triggers a pattern match. Gestalt psychologists in the early twentieth century called this reorganizing of problem elements "incubation" and the pattern match after a successful reorganization "insight."[27] (See our discussion of the "Ah, Ha" moment on p. 103.)

Critical thinking and problem solving in the ULM involve both continued search for new sensory input (new information) and restructuring and transforming available sensory input into different configurations. We might describe this using everyday language as looking at things in new ways or "brainstorming," but the basic purpose of this from the perspective of the ULM is to create a pattern match to appropriate existing knowledge. Research in expertise has confirmed that reorganization and transformation of a problem until a known solution is found is what distinguishes expert from non-expert problem solving.[28] Assuming that students have learned knowledge in school (i.e., created a chunk or procedural chain in long term memory), their ability to transfer this to the real world will depend largely on their ability to keep examining the situation (search for more input) and keep transforming the input that they have available until they find something that triggers a pattern match. In everyday terms, we call this working memory processing "thinking." That this "thinking" can require effort is something we address in the next chapter.

This type of working memory reorganization and search can be thought of as dynamically building declarative or procedural knowledge in real time. Such problem solving/critical thinking is not necessarily directed at creating a new, permanent long-term knowledge structure. It is directed at building a temporary chunk or procedural chain appropriate for accomplishing a task within the current situation and problem. Working memory is central to this dynamic knowledge creation because it "controls" what knowledge is retrieved for addressing the problem and selects which sensory input garners direct attention. Working memory processing recombines existing retrieved knowledge and sensory input until it finds an adequate resolution for the problem or accomplishes the task. We can think of this knowledge as being "constructed" in working memory to address an immediate situational need or goal. This construction may be more or less composed of existing prior knowledge or sensory input depending on how much knowledge one has been able to retrieve. Sensory inputs are likely to predominate when one knows little about the task beforehand.

Incidental Learning

The dynamically created chunks, connections, or procedural chains created during problem solving/critical thinking can leave a long-term memory trace. Because they were attended to, they will enter long-term potentiation and might result in new permanent connections in long-term memory. That is, they have the potential for becoming a part of long-term memory. This available "trace" of the dynamically created knowledge is what we commonly refer to as *incidental* learning: learning that occurs in the context of doing another task where the task goal was not explicitly to learn. If there is occasion for this particular situation to reoccur frequently in the future, then repetition will perhaps strengthen these into new declarative or procedural knowledge networks in long-term memory. If the situation and task are truly unique, nothing will ultimately be saved, but this makes sense given that nothing really needs to be saved in long-term memory if the situation is unlikely to ever happen again.

Incidental learning is often seen as a goal of "hands-on" instruction. As an instructional method, hands-on learning and other activity-based approaches share a belief that as students dynamically construct solutions to a given problem, they will extract declarative knowledge chunks and procedural knowledge chains as incidental outcomes of the activity. As we have noted, this is true. The knowledge from dynamic working memory construction is available for long-term memory storage.

There are issues when relying on incidental learning rather than more intentional direct instruction, however. The trace of the incidental declarative and procedural knowledge is likely in long-term potentiation. Unless that knowledge is repeated, permanent storage is not likely. This means that any incidental learning activity always needs to have some kind of follow-up practice or use of the incidentally learned knowledge for permanent long-term knowledge to be created.

Perhaps more problematic, when we are engaged in dynamic construction of knowledge during problem solving, our criteria for knowledge construction is getting the problem solved. From a practical standpoint in the real world, a solution just has to be adequate. It does not have to be elegant or optimal. In any real-world problem situation, there are undoubtedly many different ways that the problem could be resolved. What this means from an instructional standpoint is that, while a student may dynamically construct a solution for a hands-on problem, nothing says that this solution has to be very high quality. Studies have shown that, in many instances, students apply simple trial and error strategies to classroom problem solving activities. These get the problem solved but don't generate meaningful incidental knowledge; they just reinforce the trial and error pattern.[29]

As we have noted before, there are advantages to establishing accurate chunks and procedural chains from the start. This will require far fewer repetitions to solidify the core elements of the chunk or automatize the procedure than working from more random experiences. If the incidental learning from a hands-on activity produces a sub-optimal knowledge representation, it will be harder to move the student to a more accurate and productive knowledge structure in subsequent instruction. From the perspective of the ULM, this is why more direct instructional approaches like studying worked examples or guided practice produce better learning of problem solving skills than pure incidental learning from hands-on activity.[30]

Knowledge and Working Memory Interaction: Expanding Capacity

You may be thinking, "Well I thought that back in the working memory chapter you said that working memory only had a capacity of four chunks and even if I give you a break on chunks that they can become really big, these huge schema and procedural knowledge chains seem to imply a lot more capacity than four." This is a legitimate question. What exactly is the meaning of "capacity limit"?

The capacity limit in working memory that we talked about earlier concerns three things: (1) temporary working memory storage, (2) focused attention, and (3) processing in temporary storage. We said that working memory storage could hold about four things, which could be attended to and/or processed/connected together. So we can think of the limit of four as applying to the contents of working memory storage.

We also said that, if any of the four things in working memory storage pattern matched to something in long-term memory, the long-term memory was activated or retrieved. After our discussion of how knowledge in long-term memory is structured and connected, we can consider retrieval in more detail. It is clear from the nature of knowledge in long-term memory that pattern matching from working memory involves declarative knowledge. Whether that match concerns knowledge about the sensory input coming in or about the presence of a condition that needs action, the initial match is to declarative knowledge. Suppose you are reading a newspaper and come across the headline, "Stocks Drop." It is very likely that this headline brings to mind notions of the Stock Market, something receiving great attention at the time we wrote this book. If we were reading a paper about raising dairy cattle, this headline might refer to a decline in the number of dairy cows in herds. Of course, sillier things are possible. Imagine a Puritan dropping one of those devices into which offenders were locked to atone for their errors. Imagine a chef dropping pots of chicken broth.

When a pattern match is made to some part of the neural network of a declarative knowledge chunk, those neurons fire. As a result, the rest of the neural network of the chunk fires and this firing spreads to any other chunks connected to the original fired chunk through spreading activation. The long-term memory neurons for declarative knowledge are not in working memory's temporary storage; they are in other brain regions. So the chunk firing is not happening in working memory's storage area; it is happening in other brain storage areas. The chunk firing is not using any of working memory's temporary storage. The chunk firing also doesn't require any attention from working memory's processing. The spreading activation continues spontaneously based on long-term neuron firing potentials. In effect, once pattern matching starts one chunk firing, the rest of the activation happens inevitably.

Unitary single chunks, no matter how large, apparently can be focused on by attention as one thing. Any *other* chunks activated by spreading activation may become the focus of attention and brought into working memory, however. Although we don't know the exact way this happens in the brain yet, the entire network of chunks activated through spreading activation appears to be "available" to be focused on by attention. While attention can't focus simultaneously on all of the activated chunks, all appear to be retrievable into working memory during the time they are active. Before the model of activating long-term memory chunks within working memory slots, Ericsson and Kintsch published a paper entitled "Long-term Working Memory."[31] They did this to resolve problems with speed of search. Actual observed memory searching is much faster than would occur if each of the various pieces had to be brought into working memory, one at a time, and processed that way. Having the entire network activated through spreading activation allows for just such rapid memory searching as any part of the network is essentially immediately available.

This is important because we noted that temporary storage in working memory is very short lived. The contents of temporary storage are cycled out very quickly. The spreading activation in long-term memory extends over a longer time period. This allows continuity of "thought" even if the original element that started the pattern match is erased from temporary storage. Working memory can reinstate the activated chunk by focusing on any part of the active network. This is how we often are able to recover when our work is briefly interrupted by something that does not require much processing.

The net effect is that the amount of knowledge potentially available for processing in working memory can become quite large. The knowledge available through attention focused on a single chunk is equal to the size of an activated chunk and any chunks associated with the active chunk. This potential knowledge is expanded even more by the fact that working memory only requires one slot to pattern match and retrieve a single chunk network. Each of the remaining three slots could itself be pattern matching and retrieving from a different chunk network. Thus, the potential knowledge available to us for processing in working memory is huge, even if working memory itself is limited to only four chunks. An excellent example of this comes from research on chess masters. If we arrange a large number of chess pieces on a chess board, show the arrangement to a person who does not know how to play chess, and ask her to recall where the pieces were, she can recall about 4–7 pieces accurately. This looks very much like working memory's span limit. This recall is not affected by whether the pieces are randomly placed on the board or are in a position that would occur in an actual chess game. If we do this same study with chess masters, an interesting thing happens. If the pieces are random (i.e., have no meaning in the context of the game of chess), chess masters can recall the same 4–7 pieces as non-chess players. If the pieces are in a game position, however, the chess master can recall the placement of up to 30 pieces accurately. This happens because becoming a chess master involves learning board positions that are relevant to play. It has been estimated that chess masters may have chunks for as many as 50,000 unique arrangements of pieces. Not only can these large chunks be retrieved, but each is connected to other game position chunks reflecting potential future board positions that can follow from the current one and to procedural knowledge chains of possible move sequences.[32]

A second expansion of working memory capacity is due to automatization. Automatization is what allows us to do multiple things or as cognitive psychologists would say parallel process. If we had to actively attend to every condition–action in a procedure chain in working memory, we would rapidly run out of capacity. Remember that we can focus on only one thing at a time in working memory, and we can only hold four chunks in working memory storage, which in this case would be a total of four conditions and actions combined. So we could only do whatever single procedural chain we had focused attention on. We would literally be unable to walk and chew gum at the same time.

Like spreading activation of declarative knowledge chunks, though, working memory needs to only attend to the initial condition that triggers a procedure. As we discussed previously, once started, an automated procedural chain can execute without the need for active attention from working memory.[33] Working memory can erase the initial condition and go to the next thing. In this way, a few procedures can be started and their executions can overlap. We can be doing automated procedures while we are doing attention-focused working memory processing on something else. But it remains true that we can only do one thing that requires attention at a time. We can multitask, but only if no more than one of the tasks requires active attention.[34] That's why driving and using a cell phone at the same time is dangerous; the phone conversation may require attention at the same time that something about the driving also requires attention.

Both chunking and automaticity dramatically increase our effective working memory capacity. While individuals do differ in their basic working memory storage span and speed of working memory processing, these differences are small compared to the capacity and retrieval speeds of long-term memory knowledge. As we previously noted about chess masters, what makes them masters isn't that they somehow have a dramatically larger working memory storage span; they can't recall any more random pieces than non-chess players. What makes them masters is having access to large chunks of chess knowledge. Similarly, what makes us able to perform skilled behaviors or problem solve effectively isn't our ability to rapidly manipulate four things in working memory. Instead, it is our learned, automated procedures that cause skilled responses and retrieval of effective learned problem solutions. It is this expansion of basic working memory capacity through knowledge that allows us to supersede the underlying biological limits on our raw processing capability.

ULM Learning Principle 2: The Prior Knowledge Effect

We now can address a key aspect of the second principle of the ULM: the prior knowledge effect. Considerable research has shown that learning outcomes are heavily influenced by how much one already knows about the subject being learned. A model developed by Parkerson et al. sets the correlation between prior knowledge and achievement at 0.72.[35] A predecessor to the ULM, the interactive compensatory model of learning, sets the correlation between prior knowledge and new learning as ≥ 0.6.[36] Prior knowledge also is the primary influence on new learning in schema theories.

The prior knowledge effect results from how knowledge and working memory interact during learning. Let's consider two learning situations. First, suppose we take new sensory input into working memory. According to the first three learning rules and the memory storage rules we discussed in the working memory chapter, if we want to get this stored into long-term memory, we have to attend to the input, repeat it, and/or apply some type of processing/transformation to the input. This then has to be retained in long-term potentiation and from there be accessed enough subsequent times for a neural change in long-term memory to occur. Subsequently, that long-term memory knowledge needs to be retrieved repeatedly to strengthen into usable knowledge or it will weaken and decay. This process is obviously "a lot of work." A large number of things have to happen both immediately and over time for this learning to occur. Nuthall and Alton-Lee provide a number of examples of children's classroom learning that demonstrate these processes.[37]

Consider the following examples taken from think-alouds of undergraduate students trying to learn about a very complex computer simulation model of procedural learning.[38]

Consists of knowledge element representing the memory or known information as condition [almost verbatim reading with underlining]
Knowledge is represented as a condition [rereading]
If such and such then so and so
If part condition, then part the action [underlines]
If part is the condition and the then part is the action [while rereading]

This student is using repetition both by rereading and by making almost verbatim restatements of the information.

Now consider a second student.

Three step cycle [underlines]
First step-matching the condition clauses of rules to the active information
Ok
Second step-selecting
And third the actual executing
Matching, selecting, and executing, which is the 3-step cycle in the PI processing [while reviewing what just read]
When it is fired-what happens [during re-reading]
Have to reread this-didn't catch it the first time
When it is fired, the action statements alter the active information allowing a new set [re-reading verbatim statement]
Each one of these 3-step cycles has one time step [underlines]
Ok, the three steps -3-step cycle is a time step [while re-reading]
Sequential series until the goal is reached [while re-reading]
Got it.

Again, in this example, we can see extensive amounts of repetition and rereading of the material. This student is clearly "working" at making meaning and getting the information into memory by basically repetition, with some minimal paraphrasing.

Now let's suppose that there already is a long-term declarative knowledge chunk for the sensory input. When the input comes into working memory, it will trigger a pattern match retrieval of the chunk. If the input is already contained in the chunk (i.e., the learner already has the knowledge), the corresponding neurons will be strengthened and we are done. If only part of the new sensory input is in the chunk, the new parts will be appended to the chunk. The new part will not have the same strength as the existing chunk, but if it is subsequently repeated, it will just become part of the chunk over time. Clearly, this process involves much less work than building a new chunk from scratch. Storage is not dependent on immediate working memory processing or long-term potentiation. The new information is stored with the existing memory chunk, so it gets into long-term memory immediately by virtue of the working memory association rule. Certainly, simple association storage will need refining. As we talked about previously in considering the hierarchical structure of declarative networks, if the new information should be distinguished as something unique like dog or cat rather than an integral part of an existing chunk, like animal, further learning will be needed.

Consider the following think-aloud from an experienced cognitive psychology professor learning the same passage as the novice in the first previous example:

Knowledge element representing memory or known information processed by the cognitive system [reading aloud]
That's kind of like – um declarative knowledge, long-term memory
Ah, knowledge represented as condition – action rules, if-then [while reading]
So really, it's more semantic or propositional type stuff, if-then that's typical way of doing things
And then its got an active information element representing current information available for-from perceptual input [while reading, almost verbatim]
That's kind of like short-term memory

Note that as this person reads through the material, the incoming information triggers retrival of prior knowledge about the topic. This retrieved knowledge allows him to recognize the new material as an instance of something she already knows and move on. There is no re-reading evident.

Consider this example from another experienced psychologist learning the same material.

PI also contains an active information element representing current information available to the system . . . previous processing. . .and a working memory *where active information is contained and processing occurs . . . by comparing the conditions of rules to active information and executing the actions specified in the rule [almost verbatim 1st reading aloud]*

OK, so it's a production system.

This person has recognized what is being described as something he already knows (a production system) and has simply subsumed the new information into the chunk.

Prior knowledge also allows new learning to be integrated into a larger knowledge organization. A totally new chunk is constrained to the limits of working memory storage. Only a small amount of information can be linked together into a new chunk at one time, so new chunks are not very elaborated. If what is to be learned is more complex, the chunk will have to be built up by retrieving and attaching new elaborations, which themselves are limited in size by working memory capacity. When existing knowledge is retrieved, however, the entire network is available, including the chunk and any broader associated chunks. The new information has a considerably larger array of potential connections and associations. The connection of the new knowledge to appropriate associated knowledge is going to be more accurate and complete when the entire scope of related knowledge is available. This is why we often try to get students to relate what is being taught to what they know or to their everyday experience.

Although prior knowledge is generally facilitative for new learning, there can be problematic aspects. If the incoming knowledge generates a pattern match and our "judgment" about the new information is that it is just another instance of the same thing, we may miss important distinctions between what is new and what we already know. Also, we may apply the wrong parts of the existing prior knowledge. Research in expertise has found that beginners often group knowledge and problems based on similarities in observable sensory features, like physics problems that all have a pulley in them, rather than grouping the way experts do based on underlying physics principles (like, say Newton's second law). This impedes making the right similarity connections between new and existing knowledge.

Perhaps the most pervasive negative prior knowledge effect in education has been identified in the "misperceptions" literature. In everyday life, we build up declarative knowledge from our general experiences with the world around us. One of the key functions of school is to transmit formal symbolic knowledge in all fields, especially science. Much of this knowledge is not directly "observable" or experienced in everyday life. Instead, it is abstract like Newton's laws of motion or even more abstract as in Einstein's theory of relativity. We may, therefore, be faced with trying to teach children things that run counter to their fairly well-formed experience based prior knowledge. Studies have identified various ways that students deal with this conflict between what they know and what we teach.[39] We as educators want them to develop an overall more sophisticated knowledge network that appropriately relates their experiential knowledge with the formal scientific principles that we are teaching. The reality, however, is that students often refuse to believe the new knowledge or separate it as "school" knowledge that has no relevance to the real world or use other sub-optimal approaches.

Basic Knowledge Processes

We can summarize this chapter on knowledge into the following set of basic knowledge-related processes.

1. If knowledge in long-term memory is retrieved, its strength is increased (the repetition effect).
2. If a knowledge chunk is retrieved, all other chunks to which it is connected are retrieved and all the connections between them are strengthened (the spreading activation effect).
3. If parts of retrieved knowledge match to working memory contents, they are strengthened; if parts of retrieved knowledge do not match to working memory contents, they are weakened or inhibited (chunk building – the law of large numbers).
4. Learning personal, episodic knowledge of one's life is easy; learning semantic (non-episodic) knowledge is hard (the Fourth Rule of the ULM).
5. If an action is successful, its connection to the knowledge of the situation in which it occurred is strengthened; if an action is unsuccessful, its connection to the knowledge of the situation is weakened or inhibited (proceduralization – the practice effect).
6. If knowledge has been retrieved, new information in working memory will be connected to this knowledge (ULM Learning Principle 2: The Prior Knowledge Effect).
7. Any active knowledge in long-term memory is accessible to working memory (ULM Principle 2: Working memory capacity is increased by prior knowledge).

Notes

1. Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, *61*(3), 315–343.
2. Kandel, E. R. (2006). *In search of memory: The emergence of a new science of mind*. New York: Norton.
3. http://en.wikipedia.org/wiki/Neuron (Accessed March 22, 2009).
4. Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
5. Edelman, G. M. (1987). *Neural Darwinism. The theory of neuronal group selection*. New York, NY: Basic Books.
6. Formisano, E., De Martino, F., Bonte, M., & Goebel, R. (2008). "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science, 322*(5903), 970–973.
7. Episodic memory appears to involve the brain structure known as the hippocampus. Henry Gustav Molaison, a person known in the scientific literature as HM, had a surgical procedure to reduce seizures in which much of both sides of his hippocampus were removed. HM had no episodic memory; to him, every time you met him was the first time he met you. http://en.wikipedia.org/wiki/Hippocampus (Accessed March 22, 2009). He remembered events and facts prior to his 1953 surgery, but none thereafter. HM died on December 2, 2008, as the writing of this book was reaching final stages.
8. Loftus, E. (1974). The incredible eyewitness. *Psychology Today, 8*(7), 116–119.
9. Loftus, E., & Palmer, J. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 585–589.
10. Loftus, E. F. (1996). *Eyewitness testimony* (Paperback ed.). Cambridge, MA: Harvard University Press.

11. The area of false memories s known as confabulations. See: Shallice, T. (1999). The origin of confabulations. *Nature Neuroscience, 2*, 588–590.
12. Solomon, R. (1986). Literacy and the education of the emotions. In S. de Castell, A. Luke, & K. Egan (Eds.), *Literacy, society, and schooling: A reader* (pp. 37–58). Cambridge: Cambridge University Press.
13. Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals* . New York: D. McKay. As cited earlier, the revised taxonomy is remember, understand, apply, analyze, evaluate, and create. See Anderson, L., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Columbus: Merrill.
14. Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
15. Pavio, A. (1990). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
16. Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.
17. http://en.wikipedia.org/wiki/Plato; http://en.wikipedia.org/wiki/Aristotle (Accessed March 22, 2009).
18. http://nobelprize.org/nobel_prizes/medicine/laureates/1904/pavlov-bio.html (Accessed March 22, 2009).
19. Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
20. This is often associated with awareness or what we think of as consciousness. We don't always have to be conscious of working memory operation in controlled cognitive processing, however, so "controlled" and "conscious" are not absolutely synonymous. See Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66; Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.
21. The neurobiological mechanism underlying automaticity is not established. It may involve myelination of the neural paths activated during whatever process is being automated. For example, see Fields, R. D. (2008). White matter in learning, cognition and psychiatric disorders. *Trends in Neurosciences, 31*(7), 361–370.
22. Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *Journal of Neuroscience, 24*(47), 10702–10706.
23. Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.
24. http://www.oneplanegolfswing.com/oneplanemembers/Tour_Pros/Tiger-Woods /index.jsp (Accessed March 22, 2009).
25. Gulyas, A., Megias, M., Emri, Z., & Freund, T. (1999). Total number and ratio of excitatory and inhibitory synapses converging onto single interneurons of different types in the CA1 area of the rat hippocampus. *Journal of Neuroscience, 19*(22), 10082.
26. Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*, 32–42.
27. http://en.wikipedia.org/wiki/Gestalt_psychology (Accessed March 22, 2009); Wallas, G., & Smith, R. (1926). *Art of thought*. New York: Harcourt Yovanovich Brace.
28. Chi, M. T. H., Glaser, R., & Farr, M. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
29. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.
30. Taconis, R., Ferguson-Hessler, M. G. M., & Broekkamp, H. (2001). Teaching science problem solving: An overview of experimental work. *Journal of Research in Science Teaching, 38*(4), 442–468.

31. Ericsson, K., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245.
32. Simon, H. A., & Gilmartin, K. J. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, *5*, 29–46.
33. This is how we account for the issues raised by Ericsson & Kintsch (1995), Long-term working memory. *Psychological Review, 102*(2), 211–245.
34. For example, see Oberauer, K., & Kliegl, R. (2004). Simultaneous cognitive operations in working memory after dual-task practice. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 18.
35. Parkerson, J. A., Lomax, R. G., Schiller, D. P., & Walberg, H. J. (1984). Exploring causal models of educational achievement. *Journal of Educational Psychology*, *76*(4), 638–646.
36. Schraw, G., Brooks, D. W., & Crippen, K. J. (2005). Improving chemistry instruction using an interactive, compensatory model of learning. *Journal of Chemical Education, 82*(4), 637–640.
37. Nuthall, G., & Alton-Lee, A. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal, 32*(1), 185.
38. Shell, D. F. (1991). Effects of expertise on strategic processes in the learning of domain specific declarative knowledge (Doctoral dissertation, University of Nebraska Lincoln). *Dissertation Abstracts International*, *51*, 3679A–3680A.
39. Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*, 99–136.

Chapter 5
Motivation

In the educational and psychological literature, motivation is a rather diffuse construct. Generally, motivation is the psychological construct used to describe those things that impel and sustain us to put forth effort. If something is easy and doesn't require us to "work" on it, we don't talk much about being motivated or not. But, when we start talking about difficult things like lifting weights or running 10 miles or hanging drywall or even mental things like writing computer code for eight hours, we start to ask about what is motivating us to do these hard tasks. Broadly, motivation is the general answer to the question of "Why" we do what we do, especially why we do things that are hard to do. At the most advanced levels of expertise, we marvel at the prodigious efforts put forth by cyclist Lance Armstrong or musician Jascha Heifetz or scientist Marie Curie.

Motivation can be talked about at a variety of "levels." At the macro level, we can talk broadly about why we chose one career over another, why we enjoy "action" movies or comedies or why we would rather eat chocolate ice cream than strawberry. We can potentially talk about the motivation behind any action we might take in the short or long term. When we want to address questions of "why" at these macro levels, there are appropriate ways to frame what motivation might be behind behaviors and choices.

Although motivation can be talked about in relation to anything a person might do, in the ULM, we are concerned with what motivates someone to learn. Specifically, we are interested in what motivates someone to do those specific cognitive learning processes that we have talked about in previous chapters. Certainly, when we consider education broadly, there are important macro level motivational questions. Why students choose particular classes or majors and even why they come to school at all are important questions for educators. We can't talk meaningfully about motivating students to allocate working memory to a learning task, if students aren't in school or working on homework or doing other school related things. We will address these broader motivational issues more in Chapter 8 as we discuss how to apply the motivational principles of the ULM to the classroom and educational setting.

The ULM itself is a learning model at a more micro cognitive level. We have been discussing the processes involved in learning in relation to neuron change and

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_5,

cognitive processing in working and long-term memory. To talk about motivation at this level, we need to talk about how motivation explains "the why" of these cognitive processes. That is the focus of this chapter.

For all of the myriad studies devoted to motivation and the rather extensive array of "motivators" identified in the literature, there is surprisingly little delineation of the specifics of how motivation works and especially how it works with cognition. There has not been a comprehensive model of motivation specifying how motivation directly impacts the underlying cognitive processes of working memory and learning. The ULM provides this model of how motivation operates in conjunction with working memory. It builds on the model first put forth by two of the authors, Brooks and Shell.[1] They argued that motivation was an integral part of working memory and influenced how much of available working memory was allocated to a task. We begin with a recapitulation of their model.

Learning Principle 3: Working Memory and Motivation

In a learning task, we have to bring new information into working memory, attend to it, and process it by repetition and/or transformation. We have previously discussed what we might think of as the "absolute" capacity of working memory as about four units. We have also discussed how the effective capacity of working memory can be increased by chunking and automaticity, both of which free some of the four available storage slots and allow access to larger amounts of prior knowledge by working memory. So, working memory has available for allocation whatever is not being allocated to some other task. This is working memory's effective capacity.

We haven't dealt with the question of how much of this potentially available working memory capacity is actually being used for learning. Nothing about being "available" necessarily implies "used." Research in working memory, especially by Cowan, Engle, and their many colleagues, has distinguished between the absolute working memory capacity that a person has available and the amount of that capacity that is actually being used at any given time.[2] Ellis and Ashbrook put forth the resource allocation hypothesis: performance on any task is dependent on how much of working memory capacity is being allocated to the task rather than on how much absolute working memory capacity a person may have.[3]

If effective working memory capacity is dependent on how much of the available working memory capacity a person is allocating, then the question becomes "What determines if and how much working memory capacity a person is allocating?" This is where motivation enters the picture. In the ULM, *motivation is the primary influence on how much of available working memory capacity is actually used*. Students may have all of their working memory capacity available, but if they are not motivated to focus their attention on the learning task, and allocate their capacity to that task, they likely will not learn anything.

We have noted throughout the previous discussions of learning, working memory, and knowledge that the learning process requires effort. The cognitive actions involved in what we have described before as controlled processing in working memory are not effortless. Directing attention, repeating, creating connections, or performing transformations, all require cognitive effort. These effortful cognitions are necessary for learning all non-episodic knowledge. Learning semantic knowledge always will require effort. This is expressed in the fourth rule of learning (learning can be effortless or require effort). We can now expand this as follows: knowledge that does not require controlled processing does not require effort and hence does not require motivation; knowledge that requires controlled processing requires effort and therefore requires motivation.

The Neurobiology of “Biological” Motivation; Drives

To begin to understand how motivation and working memory interact, we start with neurobiology. The oldest studies in motivation focused on biological motivators. These studies examined what came to be generically termed *drives* and included hunger, thirst, sex, and so forth. Drives were seen as directing behavior. When you are hungry, you focus your attention on food and initiate “food seeking” activities. Extensive anatomical and physiological mapping of drives identified important areas in the hypothalamus and thalamus regions of the brain that regulate most drive related behaviors. These areas are connected through neurons to other areas important in hormonal and glandular regulation. Connections are also present to the higher cortical areas implicated in working memory including the prefrontal cortex.

Drive research likely predominated in early motivation studies because so much of the learning research of the early 20th century was dominated by animal research (such as the studies of Pavlov and Skinner already mentioned). Researchers needed ways of motivating their animals to perform experimental tasks such as running a maze or pushing a bar – things animals rarely do on their own. Since we can’t talk to animals, researchers relied on biological drives; motivating animals by making them hungry or thirsty. Recently, advanced techniques have been applied to selectively stimulate neurons in freely behaving animals “sufficient to drive behavioral conditioning.”[4]

It is true that drives motivate behavior. People who are hungry are not likely to focus much attention on anything else but food. While biological drives underlie considerable “macro” level behavior, they supply little of the motivation for learning processes in working memory. Biological drive is important, especially if these drives are not being met. A student can’t focus attention on learning if she is hungry, which is why schools have instituted free and reduced-price lunch and even breakfast programs for children who are not getting enough food at home. Teachers often ascribe some disruption in school produced by adolescents to their sexual drives. These drives, however, are not directing working memory to direct attention to learning materials being covered in a class.

Extensions of Drive Theories

The extent to which we have additional biological drives beyond those for basic survival and reproduction is unclear. One of the better-known theories that tried to extend drive-based motivation beyond basic biological needs was Maslow's Hierarchy of Needs.[5] While there was never any extensive research showing that Maslow's higher levels of needs had specific biological or neurological foundations, his hierarchy has a common sense reality that teachers can easily recognize. For example, he proposed that immediately beyond physiological drives was a need for safety (feeling secure and protected). Studies of bullying have documented that students who feel threatened or unsafe in school are not likely to be focused on learning.[6]

Maslow also noted social needs. Humans are social animals. Our existence depends on social interaction and social groups. There is evidence to suggest that inclinations toward social interaction, a social drive if you will, have biological underpinnings. Certainly school is as much a social community as an academic setting. Studies have shown that students are likely to be pursuing social goals in school as much as learning goals.[7] Also, students who are socially isolated can have negative personal and academic outcomes. As with biological drives, these needs influence behavior primarily at the macro level rather than at the level of working memory allocation in the ULM.

Beyond Drive

In the context of school learning, drive-based and need-based motivators do not play much of a direct role in student motivation. If unmet, these drives and needs can be detrimental to learning because they direct attention away from the learning task to competing thoughts or behaviors. We can make the broad argument that most learning ultimately is about drives and needs. The argument is that we learn so that we can improve our prospects of better providing for food and shelter for our future family. This connection is indirect, however. When a student allocates her working memory to studying algebra homework, she likely is not being motivated by how this assignment is going to help her eat; if she is thinking about being hungry or eating at all, it is likely to be a distraction focusing attention away from the homework.

These limits on drive-based motivations as explanations for learning and achievement have led contemporary motivation theorists and researchers to focus on two areas. The first is what we can broadly call cognitive motivators. Maslow described these as needs for the self: for recognition, achievement, esteem, respect, and confidence. Contemporary research has greatly extended our understanding of these motivators by examining them not only as drives but also as beliefs or knowledge. The second area is emotion. Both cognitive and emotional motivators influence working memory allocation.

Goals

The most basic cognitive motivator that has been studied is goals. In some of the earliest work in motivation for achievement, Atkinson noted that achievement motivation was striving for attaining a goal.[8] Extensive study of goals, summarized by Locke and Latham, led to what is known as the goal gradient hypothesis: the higher the goal that is set for performance, the higher the performance.[9] Educators will immediately see the connection to "standards."

Why are goals so closely tied to performance and achievement? Previously we have discussed controlled versus automatic processing in working memory. By definition, controlled processing implies a goal; we wouldn't be controlling things if we didn't have a purpose. In virtually all cognitive and neural net models, goals drive processing. They are what processing is moving toward. The same is true of all problem-solving models. Problem solving is about attaining a goal – the solved problem. In working memory, goals are not just motivational in the sense of increasing effort. They also allocate attention by indicating what should or shouldn't be attended to in sensory input or retrieved knowledge; they direct where attention should focus in the future. The goal sets pattern matching parameters for what is kept in working memory from sensory inputs and retrieved knowledge; it helps determine how long this information is kept in working memory storage and whether it is repeatedly attended. Goals also define whether additional transformations and connections will be made. Basically, goals are integral to all controlled processing in working memory. Drives like hunger and other macro level motivators are enacted through setting goals for working memory allocation. Working memory allocation is directed to a task like learning as long as that task is the active goal. If the goal changes, working memory allocation shifts to the new goal and begins a new task. Motivation to persist in a learning task and undertake repetitions and transformative or connective processing is dependent on the learning goal being maintained over time in the face of competing goals.

Of all motivators, goals most clearly reflect the integration of motivation and working memory proposed in the ULM. Goals motivate their attainment. Once we have a goal, we exert cognitive effort to attain that goal. This means that we direct working memory allocation to achieve the goal. A substantial amount of the motivation for controlled processing comes from the goal that is present. Although this integration of goal-directed motivation and cognitive processing in working memory is straightforward, there has been little recognition of this prior to the ULM. Cognitive theories have noted the role of goals in directing attention and cognitive processing, but have not typically considered the motivational processes involved. Contemporary goal theories have focused on the motivational properties of goals, how they influence level of effort, choice, strategy use, and achievement. They have not focused on how this motivation is expressed in working memory.[10]

Not all human thinking and behavior is goal directed. All of our working memory processing is not controlled. We can simply be "daydreaming" and taking in sensory input or retrieving knowledge as a "stream of consciousness" with no particular purpose or result in mind. *All intentional learning is goal directed, however*. With the exception of implicit learning, learning requires controlled processing. The mere acts of focusing attention on something in working memory, repeating something, or processing something to store it in long-term memory implies a purpose: a goal to engage in these working memory processes.

Although everything we speak of in this book is biological at its core, in the sense that it ultimately deals with tissues, cells, and biochemicals, goals are cognitive rather than biological entities. While drives may establish a biological imperative to, say, find food, the specific goals set for this are cognitive, based on knowledge. Through episodic experience and as a consequence of specific teaching, we have learned how to achieve important goals like locating and securing food. For early humans, this might have been where game could be found or where vegetables could be gathered. For modern humans, this might be the route to the grocery store. Our ability to establish goals is based on our knowledge of the world and how things in the world are related. As we discussed previously, goals are part of the procedural knowledge chain, often establishing the parameters for selecting one path over another and for evaluating the success of a procedure. Procedural knowledge is strengthened by goal attainment or weakened by failure to achieve a goal.

This means that goals work in conjunction with working memory the same way any declarative or procedural knowledge works. Goals can be learned. Goals can be incorporated into chunks. Goals are retrieved from long-term memory by pattern matching. Goals can be embedded into procedural knowledge chains and automatized. We retrieve goals because they are pattern matched by sensory inputs coming from situations in the environment. We also set goals based on emerging needs during problem solving/critical thinking. As situations arise that we don't have answers to, we use incoming sensory input and our knowledge to dynamically create goals in working memory, as we dynamically create knowledge. You might say to yourself something like, "*ab* didn't match with *xy*; what happens if I try to match *cd* with *xy*?" As you bumble along in your attempts, you may set and achieve or fail to achieve many, many goals. *This makes goals highly individual; even in the same surroundings, like a classroom, we cannot assume that everyone is pursuing the same or even similar goals*.

Goal Value

The motivating property of a goal is affected by its *value*. Early research recognized that it isn't just having a goal that drives goal attainment but also how much the goal is worth. This was apparent even in animal research. Changing the amount or taste of food, its "value," affected how much effort animals put forth. It wasn't just a case of being hungry; it was also a case of how much the food was "worth." [11] Goals that have higher value produce higher motivation. While value is not an absolute determinant of which goals will be selected to work on, high value goals tend to be selected over less valued goals. Value also helps sustain goal engagement. As we noted previously, goals direct and sustain working memory allocation during the time they are active. Goals with higher value are retained longer in the face of competing goals thereby sustaining working memory allocation to goal related processing longer. This is important for learning related processes like sustained attention, repetition, and transformations/connections, all of which require working memory allocation be maintained for the time necessary to complete the process.

Determining the value of a goal is not simple. Even drive goals like food do not produce a simple additive increase in motivation as the amount of food increases. More complex sensory features like smell and taste are also involved. Consider chocolate versus mashed potatoes. Small portions of chocolate may be a bigger motivator than a large serving of potatoes. On the other hand, this isn't universal. Not everyone likes chocolate, or likes it better than mashed potatoes. In humans and even in animals, value seems to be subjective. This relates to school because not all students are motivated by the same things. Some students may place little value on goals related to learning or school. Therefore, they aren't motivated by the typical goals and outcomes available in the classroom.

Why is value subjective? Like goals, value is based in our knowledge. Things are valuable not just because of their intrinsic properties, like edible food, but also because we have learned to value them. Paper money has no intrinsic value; it is a piece of paper. We have *learned* to value it because it can buy things that we do value. Other values are more personal. We value things that lead to good feelings or emotions and will discuss these later on. While we may have biologically-based mechanisms that draw us to dance and music, which are almost universally practiced and valued, we have learned to value specific aesthetic expression of these like ballet and opera.

Contingencies: The Experienced Past; The Expected Future

Contingencies reflect regular, already experienced relationships between knowledge, behaviors, and outcomes. Contingencies are the probabilities that the occurrence of one thing such as a tree is associated with the occurrence of another thing such as a leaf. You may recognize this as the basic mechanism of chunk formation and connection of chunks into large networks in declarative memory. Bits of knowledge are grouped into chunks and chunks are connected based on the frequency with which they have fired together. Stated in another way, this cumulative historical frequency reflects the contingent relationship between those bits of knowledge. The higher the cumulative frequency or contingency, the greater the strength of their neural connection.

You may also have recognized that the basic structure of procedural memory, the "If – Then" connection between conditions and actions, reflects a contingency. In our procedural knowledge example of "If it is raining, then open an umbrella," the opening of an umbrella is contingent on the presence of rain. We only do it if rain is falling.

Contingencies underlie our episodic memories. Recall from Chapter 4 that episodic memory chunks are frequency counts of the recurring commonalities of our lives. These recurring commonalities, our episodic schema or habits, reflect contingencies that have regularly occurred in our experience. For example, there is a regularity with which objects, actions, and events in our world go together. Our ability to walk into our home and recognize it as our home is contingent on the furniture, decorations, colors, smells, etc. being the same as they have been before. Our lives have consistency to the extent that things and events in our life have high contingencies; they go together and occur in predictable ways. If they don't, then we will be confused and uncertain, as we would be if all the furniture in our home were rearranged when we came back in the evening.

These episodic contingencies form the basis of a cluster of motivators known as expectancies. A contingency reflects the cumulative frequency count of how things have been connected in the past. By the law of large numbers, this cumulative frequency produces a probability that these previous connected things will occur together in the future. This probability is an expectancy. This expectancy allows us to predict how likely it is that things will occur in the future like they have occurred in the past. If I have gotten good grades when I have studied in the past, I will have episodic knowledge that grades are contingently connected to studying. The more often in my life that studying has led to high grades, the stronger the contingency between studying and high grades. This leads me to expect that, in the future, studying will pay off with high grades.

Although expectancies are anchored in our episodic memories of the contingencies we have experienced in our lives, they can also be influenced by vicarious experiences. We don't have to experience a car accident directly to learn that bad driving is contingently connected to having accidents. We could see someone else have an accident. Perhaps we could see a film about accidents or be told about what causes accidents in a defensive driving course. Therefore, a contingency can be learned just like any other kind of semantic knowledge. As a result, my expectancies are based both on my personal, episodic contingency knowledge and any other contingencies I may have learned as semantic knowledge. In fact, learned semantic contingencies may dominate for young children who have limited personal experience. Recently, evidence for neuronal changes resulting from fictive events has been found.[12]

Because contingencies can be learned, the expectancies based on them can be changed through explicit teaching and intervention. *Because contingencies are knowledge, they engage with working memory just like any declarative or procedural knowledge*. Pattern matching retrieves them. Retrieved contingencies are motivational because they produce expectancies that allow prediction of what to expect for possible alternative actions. These expectancies allow what Bandura has called forethought, the ability to anticipate future consequences of our actions and try out alternatives "in our head."[13] In ULM terms, this means trying out in working memory before implementing actions. We are motivated to pursue those actions that we expect are most likely to lead (i.e., most contingent) to achieving our goal.

Expectancies are based on multiple sources of contingency knowledge, our episodic contingencies, and any vicariously experienced or learned semantic contingency knowledge. Pattern matching in working memory will retrieve any relevant contingency knowledge chunks. Each of these specific knowledge chunks produces a probability estimate that is its expectancy. Recent animal studies provide support for the neurological basis of such probability mechanisms.[14] The actual expectancy one has, however, is the aggregate probability derived from all of the available contingency chunks. This makes expectancies "subjective." They are constructions rather than exact copies of any specific contingency. As such, they can deviate from what an observer might see as the object or true contingency. They might even deviate from our own actual past experience if we have other conflicting semantic contingency knowledge.

Expectancies motivate increased working memory allocation beyond just the neural strengths of the procedural knowledge and the goal itself. Because expectancies are based on contingencies that reflect past experiences, they bias action and working memory allocation toward those things that have worked in the past. If something has worked in the past, it is more likely to be tried again and again until it stops working. This bias toward what has worked previously has obvious adaptive utility. If we have contingent knowledge that something will work in this situation to achieve a goal, we can predict that this will likely work again. This keeps us from "starting from scratch" each time we need to decide what to do.

On the other hand, this is also how we can get stuck in a sub-optimal behavior. Although others may think and studies may support that other approaches work better, we continue using a less effective approach that has worked for us before. One of us does not "touch type" in spite of school courses on typing, admonitions from numerous colleagues and students, and a sense of envy when a touch typist reels off a passage at a keyboard in our presence. We do well with two fingers, and hold out hope for the really effective dictation computer program (If it ain't broke, don't fix it – old saying.)

Specific Motivational Expectancies

There are three contingency-based expectancies that have been found to be especially significant for motivation: means-end, outcome, and success. These three impact decisions about goals.

Means-ends expectancy is our expectancy that certain actions will lead to accomplishing a desired goal. Means-ends expectancies can be expressed as an if – then statement such as, "If I study, then I will get a high grade." In this respect, a means-ends expectancy is based on the episodic contingency record of the past success or failure of procedural knowledge in attaining various goals. Every time some procedural knowledge fires, the episodic contingency for that procedure is updated with the resulting success or failure of the procedure in achieving the goal. As the procedure is used over time, the strength of the subsequent expectancy increases. For any specific goal, there can be multiple ways to achieve it. Means-ends expectancies reflect the likelihood of success for each of these possible alternative actions and motivate us to select and pursue those actions most likely to succeed.

Outcome expectancy is based on the contingency between goals and other outcomes. Means-ends and outcome expectancies reflect two sides of a coin. Means-ends expectancies answer the question, “If I want something (e.g. a good grade), what are the possible ways that I could get this (study, bribe the teacher, cheat, pray)?” Outcome expectancies answer the question, “If I accomplish this (get a good grade), what else can I expect to get (pass the course, get into grad school, get a job)?” An outcome expectancy is the episodic contingency record of what has happened in the past when we have taken actions or achieved goals, what behaviorists called reinforcement history, and therefore what is likely to happen if we take that action or achieve that goal again. This includes the intended consequences and other possible unintended consequences that might result. They motivate us to pursue goals and take actions that lead to more valued outcomes.

Success expectancy is based on the contingency between our attempt to achieve a goal and our success in achieving it. If we have repeatedly attempted to achieve a goal but have failed, our expectancy of future success in achieving this goal is low. Success expectancies bias us to pursue goals that we have a reasonable likelihood of achieving. A good example of this is winning the gold medal in the Olympic 100-meter dash. This is clearly a valuable goal to set. There is a strong outcome expectancy that the winner of the gold medal will become rich and famous. So why aren’t we all out practicing to win the gold medal? It is because, for all but a very few, this is a goal that we don’t have very high expectancy of realizing. Considerable research in contemporary expectancy-value theory has documented the strong effects success expectancies have on effort and achievement in school.[15] Success expectancies express the overall success or failure in achieving a goal. As a result, they motivate the choice of which goal to pursue. They do not specify any particular means-ends expectancies for which actions might be relevant for successfully achieving the chosen goal.

Self-Efficacy

Bandura has noted that while outcome expectancies can tell us what we might get by achieving a goal and means-ends expectancies might tell us what actions are needed to achieve a goal, these will not be motivating if we don’t think we can perform those actions. Think about children who know that reading well will lead to desirable outcomes like getting good grades in school, but don’t believe that they can effectively read. They are unlikely to be motivated to spend much time reading. Bandura called our competence to perform actions *self-efficacy*.

In the ULM terms, which are very similar to how Bandura has described self-efficacy mechanisms, self-efficacy is our subjective probability of being able to effectively perform an action or in knowledge terms, execute a procedural knowledge action sequence.[16]

When a procedure fires, the neurons in the procedural chain are strengthened through their firing. As procedures are repeated, they increase in strength. Means-ends expectancies are based on the contingencies between the individual conditions and actions and between sequences of condition – action procedures. Self-efficacy is based on the strength of these connections. Recall that, prior to automatization, procedures are retrieved during dynamic, controlled processing in working memory. In dynamic processing or problem solving, procedures execute based on their strength. The likelihood of a procedure executing is dependent on its strength relative to other procedures. As procedures are selected in this competition, those that are selected more often increase in strength faster and have a better chance of selection next time. Self-efficacy is a knowledge chunk that keeps a trace record of this cognitive activity; an episodic memory of working memory selection. Self-efficacy, therefore, can reflect a probability from zero for a procedure that is never selected to almost 100% for an automated procedure that will execute almost every time.

As with goals and expectancies, retrieved self-efficacy knowledge enters into dynamic problem solving in working memory along with other retrieved knowledge and sensory inputs. During forethought, self-efficacy motivates working memory to attend to those procedures with the higher probabilities of success. The motivation from self-efficacy appears to be stronger than that from expectancy. This makes sense; even if a procedure would be contingently connected to achieving a goal, there isn't much reason to perform it if it is unlikely to work. Self-efficacy motivates our attention to procedures that we are likely to perform successfully.

Self-efficacy derives from numerous factors, but primarily from the past successes and failures of our procedures. As with contingencies, self-efficacy also can be influenced by vicarious learning. Seeing someone else successfully do something may increase our self-efficacy. Feedback that we can do something successfully also may enhance our self-efficacy. Self-efficacy is easier undermined than instilled, however. Suggesting that we won't be successful can lower our self-efficacy. Self-efficacy can be influenced by comparisons with others; even if we can perform an action well, we might have lower self-efficacy because we see someone else do it better. Self-efficacy is subject to our emotional state: we may be exhilarated and believe it's our time to "go for the gold;" we may be too exhausted to attempt something we normally would feel confident about accomplishing.[17] As with expectancies, self-efficacy is not an exact copy of our own experienced procedure strength. Like expectancies, it is a construction based on multiple sources including our episodic record of working memory selection and any vicariously experienced or learned semantic efficacy knowledge. Our self-efficacy at any point in time is aggregate probability derived from all of the currently retrieved efficacy knowledge. This is why Bandura refers to self-efficacy as a *subjective probability*, rather than an actual true probability of a procedures' success.[18]

Students are constantly learning new skills, such as long division, that they haven't experienced before. Many possible procedures that we may have, especially those in the process of being learned, do not have extensive histories of repetition. We may only have done them once or twice. Our experience with success or failure after only a couple of long-division problems can produce distorted self-efficacy probabilities. Also, since self-efficacy has little actual frequency to count in these instances, it may be extensively based on vicarious learning and feedback.

Self-efficacy's accuracy and its motivational strength go hand in hand with the actual strengthening of procedures through practice. As we get better, our self-efficacy increases accordingly and more accurately reflects the true probabilities of the procedure being successful. There is a general self-efficacy that does come into play in classrooms. Although a student may have little or no experience in learning skill xyz, she may be a generally successful student who already has learned abc and def. This learner tends to adopt a positive sense of self-efficacy for related tasks.

Emotion

Goals, expectancies, and self-efficacy are described in terms of cognitive motivators. Emotions also can serve as motivators. Emotions are rooted in biological systems associated with pleasure, pain, and arousal. There is a growing body of work in the neurobiology of emotion. LaBar and Cabeza have documented considerable interconnection between brain areas in the prefrontal cortex associated with working memory, and lower brain areas associated with emotion (e.g., amygdala).[19] Also, connections exist between emotion areas and the hippocampus, a primary site of long-term potentiation for episodic memory.

Emotions, sometimes also called *affect*, have multiple roles in cognition. First, they interact with working memory during ongoing cognition. Emotion is sending inputs into working memory along with the inputs from the senses and retrieved knowledge. The dynamic processing in working memory not only involves knowledge but also involves emotion. In working memory, emotion serves as a motivator: it directs working memory attention to things with positive emotions or in some cases to negative emotions such as reoccurring negative thoughts associated with depression. New emotional input can redirect attention away from a current goal. Feelings from emotional inputs can sustain continuing working memory processing allowing for repetition or transformational processing. There is evidence that emotional inputs occupy slots in working memory's temporary storage, so emotion can compete for storage space with other sensory input or retrieved knowledge.[20] In a learning situation, emotion can reduce the working memory capacity available for the learning task. This is evident when students become anxious and cannot focus attention on studying or taking a test. On the other hand, positive emotions can sustain effort and persistence in learning activities. While a student cannot allocate more working memory capacity than he has, emotions can help him to maintain near total allocation.

Emotion is intimately connected with knowledge. Emotional connections for episodic memories can be quite strong. In fact, strong emotional attachment can keep an episodic memory from decaying. Anyone reading this book is almost certain to remember what he or she was doing on the morning of September 11, 2001. Importantly, episodic emotions can interfere with declarative learning. If a student has a strong negative episodic emotional memory of school, simply entering the school building might retrieve this negative emotion. This negative emotion could take up working memory capacity and focus attention on the emotion rather than on the classroom and learning tasks. Negative emotions could produce avoidance goals that might include dropping out of school. Conversely, positive episodic emotions associated with school could provide additional motivation for learning tasks. That is why caring teachers and peer friendships are so crucial to the success of students in school, especially for those who struggle. These positive emotions can make struggling students resilient and therefore more likely to persist and succeed.

Emotional Content of Knowledge

A rather remarkable experiment was performed about the impact of odor. Two groups of students were taught in rooms pervaded with odor.[21] One group experienced camphor, a somewhat noxious chemical once used in mothballs. The other experienced chocolate. When tested, half of the camphor students were tested in the camphor odor and the other half in chocolate – and similarly for the chocolate students. In the end, camphor-taught students tested better in camphor, and chocolate-taught students tested better in chocolate. Clearly, the odors had nothing to do with the instructional content. How did this come about? Pattern matching is needed for recall (and testing). Having the same odor present gives just one more thing for a pattern to match against, and so goes the explanation.

In a similar way, emotions can be connected to declarative knowledge. Early research in what was called "mood state memory" showed that knowledge that was learned in a happy or sad emotional state was better recalled when the subject was in the same emotional state as during learning.[22] These connections to knowledge occur because emotion interacts with knowledge in a fashion similar to how any feature in sensory or proprioceptive input interacts. If an emotion is present in working memory at the same time some knowledge is present, the working memory rule of association will connect the emotion to the knowledge when this knowledge is stored in long-term memory. After that, pattern matching accounts for the seeming connection. Essentially, we create a memory of the emotional feeling that becomes part of the knowledge chunk. Emotions can afford a retrieval pattern match if an emotion is in working memory. This also works in reverse; retrieval of the chunk through a sensory match will result in the retrieval of the emotional sensation connected with the knowledge. Consider seeing a television image of a beach on Maui and remembering fondly a vacation you had with your family at that beach.

Almost all knowledge starts as episodic memory and is likely to be connected to autobiographical content such as emotions present at the time of learning. Unless emotion is integral to knowledge, emotion will not reoccur on each retrieval of the declarative content. The emotional content usually will be stripped off the knowledge chunk, as repetitions occur over time without the emotion present.

Emotions as Goals

Emotions themselves may be goals. The objective of a cognitive or behavioral action might be to produce an emotional feeling such as happiness. When educators and psychologists talk about "intrinsic" motivation or doing something for its own sake, they usually are implying an emotional goal. The activity itself produces a positive feeling; there need be no other outcome necessary. For example, *self-esteem is an emotion-based goal that often is pursued in schools.*

There is a potential conflict when pursuing an emotional goal. If a student is pursuing an intrinsic emotional goal, his working memory focus will be on achieving the emotion rather than on achieving a learning outcome. If a learning goal is also present, working memory cannot do both simultaneously; instead, it must shift back and forth between the goals. In either case, the presence of an emotional goal is likely to reduce learning because working memory allocation will not be totally on the learning task. Another way of expressing this is to say that learning is always most efficient when the learning itself is the goal. It works most efficiently when all working memory has to take into account is learning; no "distractions." Because emotions can be goals, the same cognitive motivators influence them as any other goal. Pekrun and his colleagues have found that emotional appraisals are influenced by the same value, expectancy, and self-efficacy processes we have discussed previously.[23]

The ULM and Emotion

The ULM provides a good understanding of the impact of emotion upon other components of cognition. In the ULM, *emotion acts like any other cognitive entity*. It is subject to the three principles and five rules. Like any sensory input, emotion enters working memory and, once there, is subject to the same rules for storage and connection as any other sensory input. It is stored as knowledge, and subsequently acts in working memory and cognition like any other knowledge. It is a potential component of dynamic cognition and problem solving, and has motivational properties as part of this interaction. Emotion can be a goal that, like any other goal, directs and motivates working memory allocation along with the value, means-ends expectancies, outcome expectancies, expectancy of success, and self-efficacy associated with the emotional goal (see Pekrun[24]). These interactions with working memory suggest that emotion is both a biological and a cognitive entity that in its own right can impact cognition as learned knowledge. While emotion interacts with working memory in many ways, these interactions can be understood through the principles of the ULM.

Interest

The biological and cognitive structure of interest isn't fully known. It appears to have emotional roots, as interest is typically accompanied by positive emotions. But, emotion is not always present and need not always be positive if it is. Also, like emotions, interest appears to have a knowledge-like component. We can have an episodic memory of what has been interesting and "interestingness" can be attached to a knowledge chunk like emotional memories. Excellent summaries and discussion of interest by Hidi are available.[25]

From the perspective of the ULM, interest directs attention and working memory allocation. We view interest in two ways.

Situational Interest

Situational interest is a property of sensory input from the environment. There are sensory features that appear to draw our attention. These are typically novel or unusual things. Demonstrations in chemistry and physics are thought to do this so well that many departments employ full-time demonstrators whose job is to help teachers set up and perform demonstrations. Interest appears to be associated with this novelty or uniqueness; new things are interesting. Attention to new and novel things plays a role in evolutionary selection and human development. Gibson called these affordances.[26] It is probably good to pay attention to novel things in the environment if you want to survive. Novel things appear to produce an emotional reaction that we experience as "interest." Think about the interest (curiosity) that is aroused when a new teacher joins a school. We immediately shift our attention to try and learn as much as we can about our new co-worker. As something that directs working memory attention, situational interest clearly plays a significant role in how attention is directed and working memory resources are allocated. The episodic memory of interest also would be associated with declarative knowledge chunks stored about a novel event. Interest, then, can serve in a similar way to emotion as a strengthening or motivational mechanism that facilitates retrieval. Considerable research has documented the attention-directing properties of things in the environment identified by people as interesting, and also documented the higher likelihood of recall of interesting information over other information available in a learning situation. Recent studies have identified sources of situational interest in science instruction.[27] Interest can also be a problem. As a reader tries to comprehend the important information in a piece of text, she may be sidetracked by seductive details in it. Garner and her colleagues describe these irrelevant and distracting details as ". . .novel, active, concrete, and personally involving (p. 44)."[28] They can attract a student's attention away from the more abstract and general important information that we want them to learn from the text.

Personal Interest

As people choose their profession or academic field or vocation or avocation, they become more interested in that field as they become more immersed in it and more expert. *Personal interest relates to what it is that we do*. This interest appears to be a property of experience. If we intently pursue goals, we become interested in those goals. This personal interest sustains motivation toward long-term goals. It also serves to reinforce choices that we have made for life and careers. When we lose interest, we become less motivated in these pursuits and may even shift our efforts to a new pursuit. In academic settings, this interest develops toward different subject matter and disciplines. As the situated interest from the novelty of a new subject wanes, the personal interest in the subject grows as the student engages more with it. In a sense, if situational interest arises from novelty, personal interest arises from familiarity.

We don't really know the underlying biology of personal interest. Personal interest may involve an emotional underpinning. It is most often associated with a general positive feeling, but it is not very strong. Personal interest directs attention. We are interested in the things we pursue as our vocation or avocations, and we focus attention on things related to those. At some point, some of us actually find reading technical journals interesting. Unlike other cognitive motivators such as expectancies and self-efficacy, personal interest does not obviously appear to result from a declarative knowledge chunk that is counting something specific. As we have noted, it does appear to increase in strength in conjunction with those things we most often pursue as goals. Perhaps it is an episodic memory chunk counting the number of times goals are pursued. This would be different from the counts of goal success or failure done by a contingency. Instead, it would be a sort of record of what we spent our time doing. It would then have the effect of entering working memory during dynamic processing as a motivator that essentially says, "These are the goals I have been pursuing the most, so these are probably what I should be pursuing now." If this chunk also contains a positive emotional connection, its retrieval would produce a corresponding emotional input into the decision making. Whether this is how personal interest works cognitively is speculative at this time; we have little direct evidence concerning the underlying processes involved.

Colvin devotes a chapter in his book *Talent is Overrated* to passion.[29] We would describe what he calls passion as intense personal interest. We've seen that interest in children who become experts in dinosaurs or presidents or football teams. We've seen it in quilters or bridge players or students of the American Civil War. It's a really noteworthy quality when we see it in the likes of Tiger Woods or Eric Kandel or Yo-Yo Ma. Understanding how personal interest can become so strong as to be described as passion seems to be a worthy goal.

Interest Is Idiosyncratic

Both situational and personal interests are highly individual and likely to vary greatly from person to person. This is obvious in the case of personal interest, as it arises from an individual's personal experience and choices. Situational interest also is highly individual. Although novelty has a somewhat universal character to it, what is novel is not always the same for everyone. Novelty is a function of experience; things are novel the first time we encounter them. After repeated encounters, they aren't novel anymore. As individuals have different experiences, they will experience different things as novel or not. Even when experience isn't a factor, what one person attends to as novel will not necessarily match what another attends to. The caution here is in assuming that things are universally interesting. Students will not all be interested in the same things. Some will be interested in hands-on activities; some will not. Some will be interested in reading; others won't. Some will be interested in group work; others won't. It nearly always is a mistake to assume that something is either situationally or personally interesting to everyone.

The Hierarchical Structure of Motivation

Like other declarative knowledge, cognitive motivators have a hierarchical structure from specific to general. Consider goals. We have specific goals for specific situations. These goals group at a more general level of abstraction into *goal orientations,* that is, as goals that we typically set in recurring settings. In relation to achievement motivation in school, research has identified three orientations: learning (gaining new knowledge or skill); performance (doing better than others or looking smart); and task (doing school tasks well or with minimal effort).[30] These three orientations also have been identified at even more general levels of abstraction.[31] Similarly, we have expectancies, self-efficacy, interest, and emotions about specific procedures or actions, which group, like goals, into more general clusters. In the school setting, we can talk about our self-efficacy for completing a specific assignment, our self-efficacy for an academic domain like math or reading, and our self-efficacy for school in general. General means-ends expectancies have been identified for attributing success and failure in classes, subject matter domains, and school to effort, ability, task difficulty, or luck. Outcome expectancies can be identified for specific actions, skill domains like reading or math, and for general expectancy that outcomes are contingent on behavior (called Locus of Control by Julian Rotter[32]). We can have interest or emotional reaction to specific activities, classes, teachers, subject matter (think math anxiety), or school in general. This hierarchical nesting is a function of the general hierarchical architecture of the brain and the hierarchical organization of declarative knowledge.

From a motivational standpoint, this hierarchical nesting produces motivation interactions both top-down and bottom-up. From a top-down perspective, his general self-efficacy for doing academic or schoolwork will influence a student's self-efficacy toward an individual subject like math, which will influence his self-efficacy toward a specific math assignment. From a bottom-up perspective, that student's self-efficacy for specific math problems increases or decreases her self-efficacy for math, which increases or decreases her self-efficacy for school.

Cognitive motivators have a reciprocal influence throughout the hierarchy. Influence works top-down to affect the level of a motivator at the bottom-most level of a specific action. The outcomes of that action change the immediately associated motivation, which then propagates upward through the higher levels of the hierarchy. The better you think you are, the more successful work you do; the more successful work you do, the better you think you are. As you might expect, consistent with general spreading activation, the influence of a level is less at each more distal level. Thus, school-level motivators influence subject matter motivators more so than specific activity-in-a-class motivators. In the bottom-up direction, changes in specific-activity motivators produce more change in subject matter level motivators than in school-level motivators.

This hierarchical structure allows motivation to be both dynamic and stable. Recall that learning is a process that involves the immediate, real-time allocation of working memory capacity to a specific on-going task. Working memory has limited storage; it cannot be retrieving lots of information into working memory along with the new information that is supposed to be learned. The immediate motivation for the task is being derived from the specific goals, expectancies, self-efficacy, emotions, and interest associated with the immediate task. These could be highly variable, depending on how the learning task unfolds over time. As a result, rather extreme shifts in motivation are possible. The hierarchical nesting of cognitive motivators, however, allows top-down input from more stable subject matter and school level motivation. Because each higher level knowledge chunk counts across a larger number of inputs, the probabilities in its chunk are considerably more stable than those of a lower level. Waning motivation for a task that is not going well can be compensated by higher level motivation that may say to the student, "I have succeeded in the past and can expect to succeed in this kind of task if I just stick with it." This is all a function of cognitive motivators being declarative knowledge chunks that have hierarchical organization and that can be retrieved through spreading activation.

Motivating Working Memory Allocation

Up to this point, we have discussed the various cognitive and emotion based entities that contribute to motivation, but we haven't put their effects on working memory allocation together. If we think about the vast array of sensory inputs, knowledge chunks, and possible procedural chains that could be operative at any given time, we can think of the motivators we have discussed as working like a "funnel" reducing this overwhelming number of potential working memory elements to a more manageable number. Not to just any manageable number, however, but rather to a reduced set of elements which optimize the functionality of what is selected. This funneling works as follows.

1. Goals focus our attention on only knowledge and sensory inputs relevant to the goal, eliminating anything not related to the goal from consideration.
2. Value and outcome expectancies limit goals to those most likely to produce the most valuable, important, and desired outcomes.

3. Means-ends expectancies focus attention and knowledge retrieval on the possible means we have for achieving the current goal.
4. Self-efficacy directs selection of means to those we are best able to do.
5. Positive emotions and interest sustain attention and allocation to our selected goals and means for achieving them.

Because these motivators reflect our episodic history, they channel working memory allocation to the knowledge and behaviors that have a history of working for us. As a result, they allow us to project our past experience on to our current and future working memory allocation.

Notes

1. Brooks, D. W., & Shell, D. F. (2006). Working memory, motivation, and teacher-initiated learning. *Journal of Science Education and Technology, 15*(1), 17–30.
2. Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press; Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory.* Cambridge: Cambridge University Press.
3. Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behavior* (pp. 25–43). Toronto: C. J. Hogrefe.
4. Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G. D., Bonci, A., de Lecea, L., et al. (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science, 324*(5930), 1080–1084.
5. Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370–396. For a brief review, see http://en.wikipedia.org/wiki/Maslow’s_hierarchy_of_needs (Accessed March 23, 2009).
6. Center for Mental Health in Schools at UCLA. (2008). *Conduct and Behavior Problems Related to School Aged Youth*. Los Angeles, CA: Author.
7. Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*, 202–209.
8. Atkinson, J. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.
9. Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice Hall.
10. See recent reviews in: Fryer, J. W., & Elliot, A. J. (2008). Self-regulation of achievement goal pursuit. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 53–75). New York: Erlbaum/Taylor & Francis Group; Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19*, 141–184.
11. Pereboom, A. C., & Crawford, B. M. (1958). Instrumental and competing behavior as a function of trials and reward magnitude. *Journal of Experimental Psychology, 56*, 82–85.
12. Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science, 324*(5929), 948–950.
13. Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26.
14. Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science, 324*(5928), 759–764.

15. Wigfield, A., Tonks, S., & Eccles, J. S. (2004). Expectancy value theory in cross-cultural perspective. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited: Research on sociocultural influences on motivation and learning* (Vol. 4, pp. 165–198). Greenwich, CT: Information Age Publishing.
16. Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
17. Bandura, A. (1994). Self-efficacy. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71–81). New York: Academic Press, http://www.des.emory.edu/mfp/BanEncy.html (Accessed March 20, 2009).
18. Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
19. LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54–64.
20. Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behavior* (pp. 25–43). Toronto: C. J. Hogrefe.
21. Schab, F. R. (1990). Odors and the remembrance of things past. *Journal of Experimental Psychology, 16*(4), 648–655.
22. Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99*, 229–246.
23. Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105.
24. Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*, 583–597.
25. Hidi, S., & Ainley, M. (2008). Interest and self-regulation : Relationships between two variables that influence learning. In D.H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory research, and applications* (pp. 77–109). New York: Erlbaum, Taylor & Francis Group; Hidi, S. (2006). Interest: A motivational variable with a difference. *Educational Research Review, 1*, 69–82.
26. Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin, Reprinted 1986 Erlbaum.
27. Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching, 46*(2), 147–165.
28. Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "seductive details" on macroprocessing and micro-processing in adults and children. *Cognition and Instruction, 6*, 41–57.
29. Colvin , G. (2008). *Talent is overrated: What really separates world-class performers from everybody else*. New York: Penguin Group.
30. Ng, E., & Bereiter, C. (1991). Three levels of goal orientation in learning. *Journal of the Learning Sciences, 1*, 243–271; Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100*, 105–122.
31. Spence, J. T., & Helmreich, R. L. (1983). Achievement-related motives and behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 7–74). San Francisco: Freeman.
32. Rotter J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*(1), 1–28.

Chapter 6
How the ULM Fits In

At this point we believe that we have made a case for a unified learning model around the concept of working memory wherein working memory capacity, prior knowledge, and allocation account for learning. It is extremely unlikely that any of our readers have come from this perspective. The purpose of this book is to introduce this unified learning model. Most readers will have at least some, if not much, of their training rooted in notions that use different perspectives and vocabularies. So far as we know, no other model accounts for motivation in the same way as the unified learning model. The purpose of this chapter is to try to tie the notions of the ULM and its vocabulary to other concepts, ideas, and theories with which our readers are more familiar.

Ability

Most teachers will tell you that successful learning depends upon three things: prior knowledge, ability, and motivation. What do we mean by ability? For over a decade, one notion that has pervaded the education and popular literatures is that of intelligence, a construct associated with mental ability.

While it is not our intent to review the literature on intelligence, some background is in order. Intelligence has at times been thought to be measured by a single number, the intelligence quotient (IQ).[1] Spearman was the first to identify a common factor underlying virtually all mental tests that he called the general factor or "g."[2] Generally speaking, persons with high IQs are thought to be mentally more able than those of lower IQ. This extends to the point that learning disability is commonly, and controversially, determined by measures of IQ outstripping academic achievement.[3] Quite some time back, the general factor g was separated into two distinct parts: *fluid intelligence* conceptualized as overall general mental capacity applicable to new and unfamiliar problems, and *crystallized intelligence* conceptualized as learned knowledge and skills applicable in those problems where extensive prior experience matters.[4]

As we have discussed previously, in the ULM, ability or intelligence is a function of working memory capacity and knowledge. There has been a growing consensus

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_6,

that fluid intelligence is basically working memory capacity.[5] For example, training on working memory has been found to improve scores on fluid intelligence measures.[6] In the ULM, working memory is what limits cognitive processing capability. This would be consistent with the notion of fluid intelligence as basic cognitive or mental capacity. Raw working memory capacity is approximately four slots. There certainly is individual variation in this, similar to the variation in fluid intelligence found in the general population. Memory for random numbers or letters, which is essentially a test of working memory span, is a common measure of fluid intelligence. We also see working memory as being a more fundamentally sound way to look at capacity than the older intelligence views of either g or fluid intelligence. Working memory is analyzable in more precise neurological and computational ways that go beyond the usual psychometric tests used for intelligence measurement.

The crystallized intelligence component of IQ is essentially synonymous with knowledge in the ULM. The ULM places a greater weight on knowledge (crystallized intelligence) in both ability and expert performance than is typical in most intelligence theories. As we have previously discussed, knowledge is able to actually expand functional working memory capacity. This interaction between the crystallized and fluid components of IQ is rarely discussed in the intelligence literature.

People can become highly skilled in spite of differences in raw working memory capacity or measured fluid intelligence. Given that there is important pre-selection in admittance to careers in these areas, physicians and scientists of differing ability achieve similar success in their professions. It is hard to find a nuclear physicist whom one might otherwise describe as mentally slow. Within a professional group (of physicists, physicians, electricians, etc.), fluid intelligence or raw working memory capacity matters less. Instead, within a group, what matters is knowledge. Knowledge, in turn, depends upon chunking. This has been confirmed by extensive research in expertise. Think back to our discussion of chess masters in Chapter 4. The ULM explains crystallized intelligence in terms of chunking.

When confronted with new or unusual problems or unfamiliar contexts, working memory capacity (fluid intelligence) matters more because there are no experientially developed knowledge chunks available to use. In situations where expert experience is likely, then knowledge (crystallized intelligence) – meaning available chunks – matter.

Heredity

Whether or not to include a section on heredity led to remarkably heated discussions among the authors. The literature on heritability is connected to measures of intelligence. Notice that, in our description of the ULM so far, we used neither the word intelligence nor the label IQ. The first time we mentioned either of these has been in the preceding section. Many people have asserted that intelligence is inherited, and they use as examples the case of identical twins (monozygotic twins). The correlations of IQ between twins raised apart are high, and those between twins raised together still higher. Based upon magnetic resonance imaging studies, "brain maps" of identical twins show strong similarities; those of fraternal twins show fewer similarities.[7] Whether measured psychometrically or physiologically, genetic influences are detected when data from twins is analyzed.

On the other hand, very strong environmental influences also are measured. For example, massive IQ gains have been found for several countries over the span of one generation.[8] It is hard to see how these gains could have occurred if, indeed, IQ is under strong genetic control. In addition, many researchers have challenged studies of IQ heritability. Recent meta-analysis and reviews suggest that early studies overestimated the correlations between twins IQ and that studies of adopted children may have overestimated IQ heritability.[9] As we went to press, another study of twins appeared which, once again, probably overestimates the role of heredity.[10]

A recent review of genetics of brain structure and intelligence concludes: "Nature is not democratic. Individuals' IQs vary, but the data presented in this review and elsewhere do not lead us to conclude that our intelligence is dictated solely by genes. Instead genetic interactions with the environment suggest that enriched environments will help everyone achieve their potential, but not to equality. Our potential seems largely predetermined."[11] They previously stated: "The significant influence of heredity on IQ has been misinterpreted to imply that there is little point trying to educate or be educated, or that IQ is somehow impervious to change. This is a fallacy because many environmental factors, including family rearing environments, socioeconomic status, diet, and schooling, influence IQ. As noted elsewhere (Plomin and Kosslyn, 2001),[12] gray matter volume may be correlated with intelligence partly because more intelligent individuals seek out mentally challenging activities that increase the volume of their gray matter."

The ULM is a book about learning. As noted, brains have a macro structure largely determined at birth and a micro structure determined through learning and experience. Obviously, brain macro structure is determined genetically as well as by its in utero environment. Differences in core working memory capacity, which as we have discussed, is likely what fluid intelligence is, may be largely genetic/hereditary, may be measurable, and may emerge when truly unique situations are confronted. Any aspirational boundaries they may seem to place on a given individual, however, most often can be transcended. The data are compelling that our working memory capacity increases as our knowledge increases. Whether one is "gifted" or not, in the end it is effort that pays. Testing with an IQ of 145 in the 2nd grade will not lead to a special outcome in the absence of effort. Testing with an IQ around 90 by no means indicates that top success is out of reach.[13,14]

While we live in a society that places great weight on talent as something one is fortunate enough to be born with, at least two books view this issue quite differently. In *Developing Talent in Young People*, Bloom reports the study of over one hundred individuals thought to be at the top of their professions including mathematicians and neurologists, tennis players and swimmers, and sculptors and pianists.[15] Their status was high. For example, all swimmers were Olympians, all tennis players had ranked in the top ten in the world, and all of the neurologists had received National Institute of Health Career Development Awards. The bottom line of this study was that there were no clear early signs indicating that such levels of success would inevitably emerge. Indeed, even though they had made clear early commitments, it was by no means a sure thing that all of these persons would have been identified early as prospects for the very high attainment they ultimately achieved.

A more contemporary work by Gladwell, *Outliers: The Story of Success*, identifies opportunity and legacy as being the key factors in determining the outcomes of the most successful members of society.[13] He asserts (p. 268), "To build a better world we need to replace the patchwork of lucky breaks and arbitrary advantages that today determine success, the fortunate birth dates and the happy accidents of history, with a society that provides opportunities for all." Later he says: "Superstars are products of history and community, of opportunity and legacy. Their success is not exceptional or mysterious. It is grounded in a web of advantages and inheritances, some deserved, some not, some earned, some just plain lucky, but all critical to making them who they are. The outlier, in the end, is not an outlier at all."

One way to look at hereditary differences is that they are really small but that circumstances lead to having them multiplied. Economists often speak of multiplier effects. Educators speak of "the Matthew Effect," also described as "the rich get richer."[16] In this notion, small differences are magnified when they are encouraged, thereby attracting a potential crescendo of actions including more practice, coaching, better coaching, etc. Putting hereditary arguments aside for a moment, consider the simple developmental consequences of birth date relative to some fixed date. Athletic teams often use age to separate levels of play; no one expects 6-year-olds to compete effectively against 8-year-olds, for example. Elite Canadian hockey players are far more likely to be born during the first quarter of a year (Jan-Feb-Mar) versus the last (Oct-Nov-Dec).[17] The simplest explanation for this is that, because children playing hockey are divided on the basis of birth date with January 1 being the magic day, those born earlier are better developed (bigger, stronger) than those born later in a given year, the criterion for dividing the level of competition. As a result, they attract more praise, better coaching, more opportunities, etc. The effect (in this case likely based entirely upon birth date) is real but small, but ends up being expressed dramatically in terms of reaching elite status. Changing the date to July 1 and waiting 25 years to determine the outcome could test this hypothesis. This is highly unlikely; the apparent face validity of the argument is so strong that it is unlikely to be tested. The parallel argument for heredity would be that a small genetic difference could be multiplied. One could make a similar argument for status. That is, a three-year-old for whom coaching lessons are purchased is more likely to attain elite performance than a three-year-old without such opportunity.

When will differences in core capacity matter most? When you know nothing, then having more slots to deal with whatever is available probably is very helpful. At the other end, when you know everything – when you are a well-studied expert who knows as much as the other best experts in the same way they do – then having more core capacity also may help. In between is where most of us are most of the time. That is, we are still acquiring knowledge, likely through deliberate practice, and automating that knowledge.[18]

Of the six ULM authors, one is obese and three are overweight. We all four would like to blame our genes. However, for each of us, decrease in calorie consumption and/or increase in exercise leads to weight loss. Darn!

Cognitive Development and Stages

Human development is obvious. It is especially obvious to those who have children or grandchildren. It is more obvious to elementary school teachers than college teachers. Clearly, there is biological maturation. Children grow physically, and this physical growth includes the brain and nervous system. There is evidence that working memory capacity increases during childhood and doesn't reach its full mature capacity of four slots until adolescence.[19]

From a descriptive standpoint, it is clear that children's knowledge and thinking also develop. In this area of cognitive development, Jean Piaget was among the first to note that children's thinking was qualitatively different from that of adults. Children don't just know less; they reason differently and perceive the world differently. Piaget was the first to systematically study this development of thinking and formulate what is probably the most widely and well known theory of cognitive development.

In the ULM, thinking and behavior are functions of knowledge in long-term memory. Changes in knowledge account for changes in thinking. Knowledge is gained through the learning mechanisms we have previously discussed. In the ULM, these learning mechanisms are properties of neurons in the nervous system. As such they cannot be separated from the biology of that system or the maturation of the underlying biology. As we noted in Chapter 2, and in our discussion of heredity in this chapter, the ULM views the macro structure of the brain as determined through genetics and the micro structure of the brain as determined through learning and experience. As such, the ULM would reject strong nativist and modular developmental theories of cognitive abilities that espouse strong gene driven development with little experiential input.[20] But, as discussed later in this chapter, the ULM is compatible with "weak" modular theories such as Geary's primary and secondary categories of learning.[21] The ULM also would be compatible with theories that see development of working memory capacity as a major contributor to cognitive development.[22] The work of Kandel on the neurology of learning from which we have drawn most of our neural learning mechanisms has begun focusing on how gene expression within the neuron interacts with environmental stimuli to produce the neural learning we have described previously.[23]

Teachers are probably most familiar with Piaget's theory. Piaget called knowledge structures schemes or schema, a terminology which has become standard in contemporary learning and cognitive theories for describing large integrated networks of knowledge. We have also adopted his terminology in this book to describe large declarative knowledge networks. Piaget viewed schemes as interacting with the world through assimilation and accommodation. He saw these as kept in balance by a process of what he called equilibration. The equilibration process responded to situations where existing schemes were inadequate in some way, either by not being able to recognize or classify something (e.g., assimilate it as an instance of X) or not being able to adapt itself to some situation (e.g., accommodate to the specifics of the environment). Equilibration operated to compensate for these disturbances or disequilibrium by changing the knowledge structure. Compensation was accomplished in part by what Piaget called "reflecting abstraction" or the ability to separate general knowledge from its particular content. In the ULM, this would be similar to the neural processes that strip out non-repeating situated aspects of a concept leaving the general properties. These Piagetian processes have also been applied to understanding concept formation and change, especially in science education.[24]

Piaget was a scientist and viewed knowledge growth and change as an interaction between the person/child as an active, self-regulating biological entity and the world in which the person was engaged. The ULM and other contemporary learning theories see learning in most respects in the same way that Piaget viewed "development." Some contemporary theories of concept learning have even incorporated Piagetian-like mechanisms.[25] Virtually all contemporary learning theories, including the ULM, view learning as an active, self-regulated process and recognize the influences of existing knowledge on perception, attention, and learning. Thus, the similarities between what Piaget saw as development and what contemporary theories see as learning are considerable.

We think that the ULM, more so than most other contemporary learning theories, shares Piaget's concern with understanding knowledge and intelligence as tied to the biology of a living organism. Memory in the ULM is not a "blank slate." Both working memory and long-term memory are neurological entities whose biological properties interact with and constrain what the senses take in from the environment. The ULM has replaced Piaget's mechanisms for development of advanced knowledge, with a set of mechanisms anchored in neural learning for developing chunks, procedures, and larger integrated networks of these. Overall, however, the ULM's view of intelligence as being a product of how knowledge is structured and organized is quite similar to Piaget's.

In education, the most prominent aspect of Piaget's theory has been the notion of stages. Piaget proposed that the development of thinking ability proceeded through a series of identifiable and distinct stages. Stage change was due to equilibration which produced qualitative shifts or restructuring of knowledge. This discontinuous process of restructuring was contrasted with incremental, continuous knowledge change produced by learning. Piaget proposed four stages: sensory-motor operations characteristic of infants, pre-operational, characteristic of children age 2–6, concrete operations characteristic of children age 7-adolescence, and formal operations characteristic of adolescents and adults.

The stage notion, however, has been the most misunderstood and misapplied aspect of Piaget's theory in education. Piaget proposed that his stages had a temporal sequence; they occurred successively over time, one after another, and in fixed order. He noted that, in the typical case for Western Europeans, these stages occurred at typical ages as noted previously. In applications of Piaget in education, most of the focus was placed on the age boundaries of the stages. Piaget himself never proposed that there were fixed age boundaries to his stages. He only asserted that they occurred in fixed order. Unfortunately, in schools, the age boundaries of Piaget's stages were often translated into a "readiness" boundary. It was thought that students couldn't be taught some content because they were in the wrong stage with stage defined as a specific age. Piaget certainly never saw the typical age boundaries of his stages as being constraints or indicators of readiness. He saw stages of thinking as constraining some types of learning, especially how concepts might be interpreted and understood. Whether students of any age could learn something, however, was not a function of their age; rather it was a function of their stage.

In the ULM, learning is "constrained" by prior knowledge. Prior knowledge is what allows a student to make sense of what they are learning and provides the expansions of working memory capacity needed to construct more elaborate chunks and interconnected networks. This constrains learning in much the same way that Piaget talked about a child's developmental stage constraining learning, or Vygotsky talked about the "zone of proximal development." One cannot understand fully the material that is too far beyond one's current level of understanding. In the ULM, this phenomena works through the prior knowledge effect.

In the ULM, we are concerned with knowledge growth in specific content areas and subject matter domains, especially those domains associated with formal schooling. Piaget always acknowledged that growth of knowledge and development of what we call expertise in specific subject matter was due to learning. He even recognized that children could acquire very high levels of expertise and thinking capability in specific topics or domains, so he would not have been concerned or surprised by demonstrations that high levels of thinking could be "taught" for specific content.[26] Piaget was concerned with global patterns of thinking and with the development of these global patterns. The ULM simply doesn't deal with knowledge at this global level.

In the ULM, all learning of specific subject matter knowledge is due to the mechanisms we have described previously. We take no specific position on whether stage-like change in global knowledge or thinking occurs. Work in connectionist/neural net modeling, however, has shown that global development in the Piagetian sense can be produced by computational and neurological mechanisms like those in the ULM.[27] From the perspective of the ULM, these processes, if they exist, would have little, if any, effect on learning as we have described it. We also make no claim that learning produces any type of change in a global pattern of thinking. A child learning algebra becomes more expert and skilled in algebra. She may even exhibit what Piaget called formal operational thinking in algebra. This implies nothing about her expertise in other subject matter, or whether or not she thinks formal operationally in general. Our position in the ULM is that, if stage-like developmental progressions do exist, knowledge acquisition in a content area would not be constrained by developmental stage and that developmental stage would not be a product of learning in a specific subject area.

Contemporary developmental theories have generally moved away from strong stage models as proposed by Piaget. In cognitive psychology, there is considerable convergence between current developmental theories and learning theories. There is general agreement that Piaget's equilibration mechanism probably is generally accurate but not specific enough to be a strong theory. Most contemporary developmental theories don't ascribe the same level of universality to global thinking as Piaget. Development is now viewed as both global and domain specific. Few still believe that some type of unitary "stage" characterizes all of a person's thinking across all domains. Contemporary developmental and learning theories, including the ULM, generally see knowledge as constructed by an active learner as opposed to passively received. Excellent general overviews of contemporary developmental psychology can be found in Flavel et al., "Cognitive Development" and Moshman "Adolescent Development."[28]

As with stages, we see no incompatibility between the ULM and any developmental mechanism that might be present. In the ULM, learning operates through the mechanisms we have described. There is no other influence. So learning would not be affected by a developmental mechanism, except to the extent that some developmental process could change knowledge in long-term memory independently from the learning mechanisms we have described. Such a change would influence the prior knowledge effect by altering existing knowledge. In any event, that would not change how learning happens, it would only change the prior knowledge being drawn upon. So the processes of learning described in the ULM would remain the same.

Vygotsky – ZPD; Social Construction

As noted, the ULM suggests that, *for efficient learning there is an optimal content difficulty for new material that is to be learned*. You've probably already heard of this under the label, Vygotsky's zone of proximal development.[29] What the ULM does is make the requirements for being in the zone of proximal development more explicit. That is, the material to be learned must fill or nearly fill the learner's working memory without exceeding the learner's working memory capacity. This intimately ties the "zone of proximal development" to the learners' prior knowledge, because in the ULM, prior knowledge is the dominant influence on how much working memory capacity one has.

You've probably heard the term *social constructivism*, perhaps also connected with Vygotsky's name. Vygotsky noted that certain types of knowledge were not properties of the physical environment (that is, not sensory knowledge). Some knowledge was symbolic; a product of human thought and creation. This included things like the meaning of words, moral codes and laws, our theoretical explanations for phenomena in the world, such as the laws of physics, knowledge of mathematics, stories, art, music, and literature. Vygotsky noted that this symbolic knowledge was a social construction. It was something created and made meaningful by social agreement among people. Socially constructed knowledge is not objective; its "truth" or accuracy derives from rules and relationships within the socially agreed symbol system. As a result, Vygotsky argued that symbolic knowledge based on socially constructed meanings could only be acquired through social transmission. It is not possible to acquire this knowledge from direct interaction with the physical environment as this knowledge is a property of human society, not a property of the environment.

Social interaction occurs everywhere. A considerable amount of our socially constructed knowledge is acquired from interactions in the home between parents and children and siblings, or from peers in social settings, or from other adults in everyday life. Probably everything that we come to know as "common sense" is acquired in these informal settings. At a somewhat more formal level, for most of human existence, work skills were learned "on-the-job," passed down from experienced workers or "master" craftsmen to beginners or apprentices. This on-the-job experience included not just physical skill building but gaining understanding of the "ways of thinking" about phenomena in the particular job, craft, or profession. Anyone working in a large organization is aware that each has its own unique "corporate culture." Similar socially constructed cultures of knowledge and meaning are central to all organized social groups from athletic teams to religious orders.

Recognition of the need for social transmission of symbolic, socially constructed knowledge is a key reason why formal schools were created. Our store of knowledge has increased and become more formalized into advanced symbol systems like written language, higher mathematics such as calculus, and scientific theories. These are not likely to be learned as part of everyday or even "on-the-job" social experiences. Schools were created to provide ways to better and more systematically transmit this knowledge.

The ULM does not deny that symbolic knowledge is socially constructed. The ULM, however, is a model of learning from the perspective of an individual. New knowledge and skills may be created or discovered in social groups, and that new knowledge may even be encoded into an artifact such as a book. From the perspective of the ULM, however, this socially created knowledge and skill have not been learned until they have entered the mind of an individual, be he or she a member of an originating social group that created the knowledge or someone engaging with an artifact containing that knowledge (like a book). Learning occurs when the neurons in the brain of a person change. This requires that socially constructed knowledge ultimately effect change in an individual.

In education, Vygotsky's social construction of knowledge has been used to promote particular instructional approaches such as collaborative/cooperative learning, scaffolding, mentoring, and apprenticeship. Whether this is actually justified is problematical. Because knowledge might have originally been socially created does not imply that it is best taught through "social" learning or instructional methods like collaborative group learning. Neither does the common informal transmission of everyday common sense knowledge or existence of "on-the-job" training imply that these are the best ways to transmit this knowledge. Vygotsky himself did not privilege instructional methods that mirrored either the social knowledge construction process or indirect transmission of everyday knowledge. He provides many examples of direct instruction by teachers including lectures as ways to teach and transmit socially constructed knowledge in schools.[30]

We would concur with Vygotsky concerning instructional methods. As we will discuss in subsequent chapters, many instructional approaches are consistent with the ULM, including the collaborative learning, scaffolding, and apprenticeship approaches commonly associated with social constructivist approaches. The ULM also recognizes that students can interact together and interact with teachers in ways that socially construct new knowledge and understanding. Learning may occur in a social context and be helped by teachers, peers, and others. Whether any of these methods are good or bad is not a property of the method, it is a property of how it is used.

Some final comments about the ULM and constructivism. Constructivism is a term associated with Piaget[31] and especially Vygotsky.[32] At its core, two things describe the ULM. First, learning involves rewiring within the learner; no wiring for knowledge pre-exists only to be revealed by some developmental, environmental, or other means. In that sense, all knowledge (neuronal rewiring) that we have discussed in the previous chapters is constructed. It is always the result of the transformations and connections built in working memory. In Chapter 4, we even noted the especially constructive processes that occur during problem solving and critical thinking that can lead to unique new knowledge constructions. The ULM is "constructivist" because the neural connections in the brain are based on plasticity that constructs the unique connections between neurons. Second, like other constructivist theories, the rewiring that we undergo as human learners with respect to acquiring the symbolic socially constructed knowledge taught in schools requires our active participation. It is true that our autobiographical lives experience effortless, if fragile, recall of our daily experiences. It also is true that, under special circumstances, we can invent environments in which important learning is achieved incidentally, say, through playing games in which the player must repeatedly attend to and apply rules that otherwise would be part of a more traditional school curriculum. In the end, most learning that society expects from schools requires effort on the part of the students.

Short-Term Memory

The notion of short-term memory predates that of working memory. Short-term memory referred to a temporary storage area where information was held for a brief period of time as contrasted with permanent storage in long-term memory. Working memory models have incorporated the notion of a temporary or short-term storage area or areas.[33] This is true of the ULM, as we defined working memory in Chapter 3 as containing a working memory storage area for temporarily and briefly holding elements of sensory input and/or knowledge retrieved from long-term memory. There has been debate as to whether short-term memory should be retained as a distinct construct separate from working memory. Engle, an early advocate of this distinction recently has revaluated the basis of this position.[34] Like other working memory models, the ULM integrates a consideration of the temporary and limited capacity of "short-term" memory into the broader operation of working memory. More technical debates about whether these are cognitively or neurologically unique "memories" do not affect how we describe working memory as operating in the ULM.

Cognitive Load

Cognitive load theory, originally developed by John Sweller, is perhaps the closest current theory of learning to the ULM.[35] Like the ULM, cognitive load theory is grounded in the idea that human capacity is limited. Johnstone applied the notion of limited capacity to science education two decades ago.[36] Cognitive load theory lacks the ULM's explicit connection to specific working memory processes, but it shares the basic idea of the first principle of the ULM that learning is a function of how limited capacity is allocated. Cognitive load theory has also recognized the second principle of the ULM that knowledge affects capacity. Where it has diverged, has been in the overt consideration of the role of motivation as an influence on capacity that is a core principle of the ULM.

Cognitive load theory has focused most extensively on how aspects of the instructional setting and materials affect limited capacity. Sweller noted that the instructional setting placed various demands on the learners' available cognitive capacity that he referred to as load. In early work, two kinds of cognitive load were envisioned. The first involved load that resulted from the inherent difficulty of the to-be-learned material called *intrinsic load*. Intrinsic load was a function of how complex or difficult the material was and how much information needed to be stored in working memory and processed. Think of the difference between learning a specific fact and learning how to balance a chemical reaction. In ULM terms, learning through the second rule by repetition would have less "load" than learning through the third rule by making connections. The second, *extrinsic load*, consisted of features of the learning environment that were not themselves part of what was to be learned, but demanded use of cognitive capacity. Typically, extrinsic load was encountered as the result of poor instructional design or distracting features in the design. Extrinsic load decreases learning because the capacity allocated to deal with it decreases the capacity available for learning. For example, putting everything on one page or one screen leads to better overall outcomes than having the learner move back-and-forth between two pages, or two screens, or a page and a screen. The ULM takes a similar view that anything that reduces allocation of working memory capacity to the learning task will decrease learning.

In more recent work, a third type of load called *germane load* has been identified. Germane load is a companion to intrinsic load and reflects a distinction similar to that in the ULM between storage and processing. Intrinsic load is now viewed as a storage load consisting of the number of unique things that must be learned and therefore held in working memory, and germane load is the load from needing to process the stored information, such as making connections or transformation.[37] As in the ULM, working memory capacity needs to balance space for both storage and processing. Consider the model of slots. Let's say you have one slot filled with knowledge about angles and another filled with knowledge about triangles, two entities creating intrinsic load, and you need to pull them together – say to create the notion that the sum of the angles in a triangle is 180°. It is in this third slot that you are going to need to handle the germane load and start tying together those sets of notions.

Perhaps the single most important overall difference between cognitive load theory and the ULM is that the former remains almost universally silent with respect to learner motivation. Sweller's most recent book shows motivation as a single index topic that directs readers to a section entitled "Emotional vs. Cognitive Sources of Motivation."[38] That section points to two studies by Mayer suggesting that the inclusion of "seductive details" in learning materials negatively impacts learning.[39]

It is extremely unlikely that the ULM would have emerged from us had we not been very familiar with the work of Sweller and others in the area of cognitive load.[40] Virtually all of the instructional principles and suggestions derived by cognitive load researchers are compatible with the ULM. We see no reason to discuss this work in detail in this book. Readers are encouraged to see Sweller's book, *Efficiency in Instruction*.[41] Very recently, an entire edition of *Educational Psychology Review* was devoted to an exploration of the frontiers of cognitive load theory.[42]

"Ah, Ha" Moments Involve Special Marking for Later Retrieval

We've all had "Ah, ha" moments (sometimes attributed to Archimedes upon figuring out how to measure volume by displacement and called "Eureka" or "I have found it" moments).[43] There is the moment when you have just finished studying and realize that you are "ready for the test." There is the moment when you finish cleaning up after painting a room and can say to yourself, "That's done, now I can do something that is fun." For the most part, when teachers speak of "Ah, ha" moments, these examples are not the ones in question.

There is copious anecdotal support for special events in learning called "Ah, ha" moments when, by obvious emoting of one type or another, the learner says, "Now I get it." In the literature, especially the conceptual change literature, "Ah, ha" moments have been viewed as reflecting qualitative shifts in thinking that result from restructuring of knowledge, similar to what Piaget proposed about cognitive development. Restructuring theories have generally proposed that qualitative shifts in knowledge result from a specific "restructuring" learning mechanism that produces qualitative change distinct from the presumably continuous change produced by regular learning mechanism, although these mechanisms have not been well defined in the previous literature.[44]

As the ULM was being developed, we at first attributed "Ah, ha" moments to routine chunking.[45] This was based partially on work in cognitive modeling that has failed to find qualitative restructuring mechanisms.[46] We now believe that "Ah, ha" moments result from something different. Back in the knowledge chapter, we talked about how one of the key processes of working memory is to connect input in working memory storage together in different ways. We noted that these new connections, which arise during episodes of problem solving or critical thinking, create or construct new knowledge. Clearly, any new knowledge arising from one of these constructions would be "qualitatively" different from the way that knowledge was interconnected previously. In the ULM, repetition is the quantitative learning mechanism and connection is the "qualitative." We would hold that the ULM's third principle of learning is the "restructuring" mechanism alluded to in the previous literature.

When we become aware of newly constructed knowledge, we can experience an "Ah, ha." As we have noted previously, working memory processes, including creating new connections, do not have to be conscious. "Ah ha" moments especially seem to emerge after an initially unsuccessful attempt to consciously construct a solution that is subsequently pursued unconsciously. These resolutions are greeted with glee when we are finally aware of the resolution. One might ask, "Why this moment of glee?" We have previously discussed the role of emotion in marking knowledge for storage and enhancing retrieval. Our experience of emotion during an "Ah ha" may be a way of marking the newly constructed knowledge to insure storage in a way to enhance later retrieval.

There is experimental neurological evidence for these insightful moments. For example, when learners are presented with triplets of words such as date/alley/fold and asked to discover a word linking them – in this case "blind." Using both functional magnetic resonance imaging (fMRI) and electroencephalogram (EEG) measurements, sudden bursts appear at the instant that the solution emerges.[47] Subjects are asked to indicate whether they solved the problem with or without insight. Insight is described as coming to an "impasse" in solving and feels as if he/she has moved on followed by a moment in which the solution suddenly emerges. Indeed, neural regions not involved in the routine problem solving are indicated for the insightful solutions.

Ordinary Learning Moments Require No Special Marking

There is an aspect about "Ah, ha" events that should be emphasized. They are typically reported as happening after an unsuccessful search. Generally, we report them after we think we have stopped consciously searching. So, the search that was successful must have been a subconscious search. With a model like the ULM, you will be tempted to think that you can control four slots and that you can consciously load them. That's not how it works. The allocation process is only partially within our conscious control. The details of most of the processes we use remain hidden from us even as we use them.

Are “Ah, ha” moments good or bad? There certainly have been some of these moments during the development of the ULM. For example, there was the emergence of the notion that motivation amounts to working memory allocation. This was followed years later by the realization that neuroscientists speak of the top-spot in working memory using the term focus of attention, and thus all of the work they had done in terms of focusing attention could then be revisited from the perspective of motivation.[48]

It certainly is not a sure thing, however, that the good feelings that come from the “Ah, ha” moment are worth the inefficiencies connected with designing instruction based upon discovery and “Ah, ha” moments. We can ask students to “think” about something or design problem solving activities that require unique applications of knowledge, but we can’t guarantee that a student will actually be able to create meaningful new connections. We once had a conversation with a college classroom teacher who developed “learning cycle” laboratory activities. It was considered to be a bad activity when all students “got it” immediately. It was worse, however, when no students got it.[49]

Savants: Prodigies

During the earliest stages of the development of this book, one author said to another, “But what about savants?” At that time, both of us held the folklore-entrenched view that savants were very unique, but important, special cases where knowledge seems to appear without having been learned in a traditional way.

Persons with savant syndrome show generally low skills and usually have problems functioning but display strong and unusual skills in unique areas (e.g., recalling day or weather given date, piano, painting). As noted, it once was thought that these skills emerged spontaneously giving support to the claim that humans are somehow pre-wired and that savants are able to tap into that wiring. According to the ULM, *no human is prewired with an advanced skill*. That is, the ULM insists that none of us can spontaneously drive or play a piano; these are skills that *must* be learned. Whatever pre-wiring humans may have, the ULM sees that pre-wiring as being rather minimal. More recent and thorough study suggests that savant skills are an outcome of much more conventional instruction: practice with feedback. A considered current view of the phenomenon is summed up as: “A person with normal brain function could become a calendar prodigy; it is just very unlikely that they will find the process rewarding enough to persevere to this extent.”[50]

The reason this is important to note is that, if just one savant were found with advanced skills that appeared in situ without being learned, then the ULM could not be correct as a model. In the ULM, microstructure is constructed as opposed to being “revealed” from some otherwise, hereditary, pre-existing states. Indeed, it seems that special skills only emerge through practice. Proportional reasoning as a general skill emerges in students as the result of practice with phenomena and materials in which this concept applies, for example.

There are those regarded as prodigious geniuses such as the musician Mozart and the chess master Fisher. Careful studies of the lives of prodigies always seem to show external encouragement, access to effective coaching, and materials such as a library. While it was said that Mozart wrote music in singular bursts (i.e., put directly from "mind" onto paper without editing), the evidence is that he edited over long time periods in a fashion similar to that of most composers.[51] Much literature points to the notion that, in order to develop expertise, about ten years of study must be completed.[52] The so-called 10-year rule appears to have applied to Mozart and Fisher as well as to other elite experts.

Special Memory

There are reports of some truly unusual people with respect to memory. For example, AJ remembers essentially all of her life experiences.[53] "Her memory is 'nonstop, uncontrollable, and automatic.' AJ spends an excessive amount of time recalling her personal past with considerable accuracy and reliability." Some suspect that she spends a great deal of time rehearsing as we would when trying to achieve rote memorization. In the absence of considerable effort, ordinary episodic memory is a fragile and fleeting thing for most of us. Rather than having this unusual ability to recall as a gift, AJ reports herself to be dysfunctional and seeks help. Other reports of really good memories, variously called eidetic memory or photographic memory, appear frequently but most often are surrounded by controversy.[54]

Experts often acquire very powerful memory. They can attend a seminar and recall nearly all of what was said, although not verbatim. This is an example of just how important prior knowledge and motivation are. Most likely, they already know much of what is said and think so often about that which is new that they are really rehearsing that content over and over. Those are characteristics of many experts and how they deal with information. Two of the authors are thought to have good memories. One, DS, is able to recall a wide range of content with some but not much effort. The other, DWB, was known for being able to call 750 of 1,000 students from lecture class by face on campus but outside of his classroom. Rather than the result of a special gift, this recognition resulted from hour upon hour of study using photographs arranged in grids that matched student's places in laboratory. It included immediate rehearsals in the teaching laboratory followed by engaging the students in conversation – something that teachers in much smaller classes but without laboratory sessions have no opportunity to practice.

Multiple Intelligences

A very popular notion developed by Gardner and rooted in studies of geniuses, savants and aphasics is multiple intelligences.[55] In this notion, intelligences may be divided into subcategories (logical-mathematical, spatial, musical, interpersonal, etc.) and individuals may essentially have more of one than another. No agreed upon measures of these sub-categories have emerged that are not well-predicted by measures of fluid intelligence, however.[56]

At the same time that no serious science has emerged to support multiple intelligences, this remains a notion warmly embraced by classroom teachers. After all, it stands the test of time: we all know people who are good at some things but bad at others. The ULM accounts for these differences in terms of prior knowledge. Further, as we become good at things, we usually become good at becoming better at those same things. We develop what Bandura calls virtuous cycles.[57]

We've cited some important evidence about expertise developed by Bloom and his colleagues and published in *Developing Talent in Young People*.[16] They define talent as " . . . an unusually high level of demonstrated ability, achievement, or skill in some special field of study or interest."[58] They go on to assert: "Although we cannot be certain of this, we believe that only a small percentage (10% or less) of these talented individuals had progressed far enough by age eleven or twelve for anyone to make confident predictions that these would be among the top 25 in the talent field by the ages of twenty to thirty. . . . Even in retrospect, we do not believe that perfecting of aptitude tests or other predictive instruments would enable us or other workers in the field to predict high-level potential talent at these early ages."[59]

Since Gardner is generally regarded as the principal advocate of multiple intelligences, it is appropriate to consider his writings in more detail. He describes multiple intelligences in terms of "biopsychological potential," cites genetic proclivity for diseases, extrapolates this heritability to intelligence, and then emphasizes assumptions of many scientists that intelligence is heritable by as high as 80% based on IQ tests.[60] While Gardner does not deny the possible influences of environment on the development of intelligence, his theory does not paint a promising picture for those who would exceed their "potential." The contrast between these views and the ULM are many. The ULM interprets learning in terms of neuronal changes: that neurons are neurons (not music neurons or aesthetic neurons), and that prior learning greatly impacts one's ability to work as what Bloom might call a "talented" person.

In the end, this probably doesn't matter. In *Creating Minds*, Gardner identifies a small sample of highly recognized individuals (e.g., Einstein, Gandhi) in which he cites the so-called 10-year rule. (The 10-year rule, generally regarded as accepted, is that it takes about 10 years of serious study with appropriate teaching, mentoring, and support to achieve expert status in virtually all fields.)[61] Gladwell calls this the 10,000 h rule.[62] The amount of time spent practicing distinguishes excellent from good musicians.[63] Thus, even if one believes that there are multiple intelligences, this in no way alters the amount of time and effort required to develop expertise.

One thing in Gardner's writing is especially noteworthy, namely, that people at the top of their talent areas frame experiences in a way that is positive.[64] Perhaps it can be summed up as the ability to learn from one's mistakes.

Learning Styles

A learning style is thought of as some combination of characteristics (cognitive, affective, psychological) related to how one interacts with one's learning environment.[65] This idea is rooted in the belief that humans differ in the sense modality of stimuli from which they learn best – take in, remember, and process new information.[66] Many studies have failed to find evidence for such styles. For example, Krätzig and Arbuthnott conclude: "The present results suggest that people's intuitions about their learning styles may be incorrectly attributed. Specifically, such styles may indicate preferences and motivations rather than inherent efficiency at taking in and recalling information through specific sensory modalities."[67] Coffield et al. present a critical review.[68] What is the take of the ULM on learning styles? First, convincing data for the existence of learning styles is not at hand. Based upon the ULM fifth rule that "learning is learning," we suspect that there are no biological constraints placed on people – that nearly any of us could end up emphasizing nearly any style (if such a thing really exists).[69] In short, we are skeptical of the existence of this phenomenon in a meaningful, measurable way. If such a phenomenon does exist, it certainly can be ascribed to learning the "style" (i.e., prior knowledge) and need not invoke some special inherited uniqueness (as in one of the growing list of multiple intelligences).

It is certainly true, however, that students think of themselves as having different learning styles. Whether objectively true or not, students can be heard saying things like "I learn better when I can hear about it in class than reading about it," or "I can understand it better when I read about it." If students think of themselves as learning better with certain content or media than others, then given what we have previously discussed about expectancies, it is reasonable to think that students develop particular motivational preferences toward different ways of learning. It is probably more proper to talk about "motivational" style than learning style. Given what we discussed in the motivation chapter, it is easy to see that if a student feels more confident or expects to learn better with visual material than through listening, he/she is likely to have more motivation and more effective allocation of working memory with visual materials.

The Executive

The notion of an executive probably evolved from tradition, possibly the tradition of a homunculus, rather than from a body of data that required explanation.[70] If you search this book, you'll not find the term executive used in the sense of CEO or "decider." The original model of Baddeley and Hitch (1974) had three parts: a central executive, a phonological loop, and a visuospatial sketchpad.[71] In 2000, Baddeley found the need to add an "episodic buffer."[72] The ULM does not make use of any of these. Instead, anything in memory is potential fodder for working memory. If this is the case, how are decisions made? In some ways, these arguments go back to the homunculus arguments: “One may explain (human) vision by noting that light from the outside world forms an image on the retinas in the eyes and something (or someone) in the brain looks at these images as if they are images on a movie screen. . . .”[73] That voice that we hear speaking to us as we decide things fools us over and over. Because of the way neural networks work, we don’t need any such “decider” in the ULM.[74]

As we discussed previously in Chapters 3 and 4, no special inductive or decision making neural or cognitive mechanisms are required to explain how working memory operates or how knowledge is stored. Attention, repetition, pattern matching, and the other mechanisms of the ULM operate without the need for a supervisor or any oversight system. While it is proper to talk about a whole person as making decisions, self-regulating, or controlling their working memory processing, this is different from the notion of a cognitive or “inside the head” executive system. We have noted that the working memory, long-term memory, and motivational cognitive and neurological processes of the ULM do not require controlled processing by the person. They can work automatically. Similarly, they do not require a cognitive executive to initiate or monitor them.

Gender Differences

There do seem to be differences between men and women in terms of cognition. Generally speaking, these differences are small. Several decades ago, gender differences in, say, mathematics scores were large and explanations for those differences many. The title of a recent *Science* Education Forum says it all: “Gender Similarities Characterize Math Performance.”[75] We do not doubt for a moment that gender differences exist and that they may be important. As with nearly everything else we assert in the ULM, however, we come down on the largest factor in creating these differences as being differences in knowledge (i.e., prior learning).[76] There are gender differences in reported self-efficacy that we discuss later.[77]

Primary Versus Secondary Learning

A special issue of the Educational Psychologist entitled “Evolution of the Educated Species” appeared late in 2008 (Volume 43, Number 4). The central paper by Geary, based on earlier work, incorporates many novel ideas.[78] A key notion is that learning can be divided into biologically primary and biologically secondary categories. Biologically primary categories refer to learning associated with specialized neurological processing areas in the brain that have evolved to optimize speed of learning and processing. We have mentioned some of these types of areas previously, like the specialized sensory processing areas for visual, auditory, and other sensory input and the motor cortex areas specialized to specific peripheral muscle groups like the fingers. Biologically primary categories can also exist for higher functions like language or at least oral language. Geary notes that these specialized areas have evolved to be sensitive to specific learning inputs related to their specialty processing.

The biologically secondary category refers to learning where no specialized neural area exists to process that knowledge. This would include basically all of what Vygotsky called socially constructed and transmitted knowledge which includes pretty much everything we teach in school. There simply aren't specialized brain areas for chemistry, or social studies, or algebra, or literature. Learning about these occurs in the brain's general non-specialized areas or as Geary has noted, by piggybacking on and repurposing one of the biologically primary areas.

A classic example of this distinction is language. Oral language is a biological primary. It is learned early in life, during infancy. Being in a language using environment seems to be sufficient for learning. Extensive special motivation does not seem to be required; in fact, motivation seems to be intrinsic. There are underlying specialized language areas in the brain that facilitate this processing through specialized functioning. Learning written language, reading and writing, however, is a biologically secondary category. There is no specialized brain area for written language; written language appropriates the oral language areas and repurposes them. Learning usually requires explicit instruction and deliberate motivation. Learning is not certain. Virtually all humans become proficient in their native oral language. Reading and writing are far less universal. Even where most of the population is literate, there are wide variances in proficiency.

Geary's arguments are an extension of older arguments about heredity and nature versus nurture that we have discussed previously. Geary's categories don't really affect the premises of the ULM. Learning in both biological primary and biological secondary areas follows the same neurological learning principles we have described previously. Neurons are strengthened in the same ways whether in a specialized or general area of the brain. Because biologically primary areas are sensitive only to certain kinds of stimuli, their frequency counts achieve accuracy faster. Geary argues that attention is drawn to sensory inputs associated with biologically primary categories. We also appear to be naturally drawn to or motivated to learn biologically primary knowledge. He argues that much of everyday learning, what we might call common sense or general knowledge, is associated with biologically primary categories that have evolved over time to reflect the common aspects of the human physical and social environment.

Geary also argues that school learning deals primarily with knowledge that is biologically secondary. The subject matter being learned in school is not going to get a boost of efficiency or motivation from specialized neurology. It will require all of the ULM rules we have discussed. This distinction also has instructional implications. A number of situated cognition, apprenticeship, and constructivist instructional approaches have tried to use everyday learning as a model. They argue that if we could make school more like everyday experience, or instruction in school more like the informal approaches used in everyday settings, students would learn more easily and be more motivated. Geary argues that this is not likely to be the case. Everyday learning is different not because it occurs outside school, but because it is biologically primary. This biological primary learning will not transfer to the biological secondary learning required for school subjects.

History and Background

The ULM did not emerge spontaneously; the search for a unified model is not new. Allen Newell gave a series of lectures at Harvard in 1987 and wrote *Unified Theories of Cognition* based on them.[79] He suggests that a unified theory must offer three advantages: explanation, prediction, and prescription for control. He indicates that a theory of cognition must deal with a range of issues from problem solving through language to daydreaming. The model he advanced was based upon his work with *SOAR*, an architecture for general intelligence first substantiated as an artificial intelligence computer system.[80] So is the ULM a unified theory of cognition in Newell's sense? The answer to this is, no. While the ULM does well in terms of explanation, prediction, and control prescription, it certainly is far from being instantiated as an operating artificially intelligent computer architecture. However, we did not set out to develop a fully realized model of cognition. We set out to create a synthesis that could unify what is known about learning, particularly classroom learning, into a workable model that fully described the process of learning.

It is remarkable how close others have been to describing what essentially is the ULM. For example, Bereiter wrote (shown below with his citations removed)[81]:

> "A contextual module is not just related to a context. It embodies the person's whole relationship to that context. The influence of culture, that initially may have been identifiable in particular beliefs, goals, and rules of conduct, can now be only globally assessed on the module as a whole. There is no longer a separate representation of the context. Instead, that representation is implicit in the whole structure of person-environment relations embodied in the module. This is what situated cognition would mean in a theory based on contextual modules. It would be an emergent property of modularity rather than an attribute of the learning process itself." [p. 613–614]
>
> "There seems no reason to suppose that any special process goes on in the development of contextual modules that is different from the kinds of processes dealt with in theories of learning and development. Computer-implemented theories have demonstrated a number of processes that could contribute to modularization. Chunking forms larger units from mental contents that are repeatedly activated together. Chunks could form across as well as within components, for instance, combining the representation of a situation with its associated affect and with goals typically pursued in that situation. Reduction to results stores the results of procedures as declarative knowledge, thus eliminating the need for the procedure; one does not need to keep solving the same problems. Finally, compilation produces streamlined procedures that eliminate calls for explicit declarative knowledge or representations of situations, because these become implicit in the procedures. What is novel in the idea of contextual modules is not the processes involved but the result. The result is a cognitive unit having properties of real human significance that are not possessed by smaller psychological units." [p. 614]

Based on this, it is clear that Bereiter anticipated chunks that were really complete up to the point of including their own motivation. Two extremely highly regarded researchers, Ericsson and Kintsch, struggled with the limited capacity of working memory and constructed the notion of "long-term working memory."[82] They expressed their dilemma as:

> On the basis of a century of laboratory research on memory many theorists have concluded that LTM can meet neither the criteria of speed and reliability for storage nor those for retrieval.

Schraw realized that ability, prior knowledge, and motivation all were important in understanding successful learning, but his model treated them separately and spoke of how one aspect could "compensate" for another without appreciating the biological underpinnings that made this possible.[83]

> We assume that each of the modules [knowledge, ability, motivation] contributes directly or indirectly to learning, and compensates for potential deficits in other components. Specifically, we assume that cognitive ability is related to learning both directly and indirectly via knowledge and regulation. Strategies and metacognition typically co-develop and are strongly related. Knowledge and regulation are related to motivation. Cognitive ability is not related to motivation. Knowledge, regulation, and motivation each are related directly to learning.

Mindset matters. How could researchers, for nearly two decades, be so close to creating a unified model? Perhaps this is because of the prior learning we bring to issues. For example, the Baddeley and Hitch paper on working memory, while seminal, included notions of separate buffers and an executive.[84] In fact, in order to account for some data, Baddeley added the notion of an "episodic buffer" as recently as 2000.[85]

We certainly are not the first to call for better integration of understandings about motivation with those about cognition. Such eminent researchers as Simon and Pintrich both wrote extensively about this need, but never developed models at the same level of detail as the ULM to achieve such integration.[86]

The biggest challenge we faced in developing the ULM was in putting parts of the old explanations and theories aside while remaining faithful to accounting for the data that led to those models.

Our Purpose in This Chapter

We obviously believe that the ULM is a powerful model that accounts for essentially all that is known about learning. The purpose of this chapter, then, has been to account for the data that others cite when arriving at their models without using those models but using the ULM, instead. So, ability becomes a mix of working memory capacity and prior knowledge rather than fluid and crystallized intelligence. Apparent jumps in skills become chunks rather than stages and are accounted for in the ULM in terms of prior knowledge. Matching learners to instructional materials becomes a question of working memory capacity and prior knowledge rather than zone of proximal development. Knowing people with vastly different skill sets becomes prior knowledge rather than multiple intelligences or learning styles. Those neurons in the cerebral cortex "learn" things in the same way no matter which cortical tissue mass or lobe they are in. "Ah, ha" moments arise when chunks are activated at the same time in working memory, although they rarely if ever have been activated at the same time before.

In a few but important cases we depend upon better data. Savantism emerges as an example of knowledge resulting from learning rather than special gifts or wiring, a view that contrasts with part of the generally accepted lore of learning.

The ULM has something going for it that scientists always strive for: parsimony. That is, we make fewer assumptions than nearly any other models. We assert that learning depends upon your working memory: how much you've got of it (capacity), how you've used it in your past (prior knowledge), and how you are using it now (motivation). The ULM not only accounts for extant data, but it implies both ways to test the model and to apply the model.

Notes

1. http://en.wikipedia.org/wiki/IQ (Accessed March 23, 2009).
2. Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.
3. Harrison, P. L., & Flanagan, D. P. (2005). *Contemporary intellectual assessment: Theories, tests, and issues*. New York: Guilford Press.
4. Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253–270.
5. Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence, 32*, 277–296; Kyllonen, P. C. (1996). Is working memory capacity Spearman's g? In I. Dennis & P. Tapsfield (Eds.), *Human abilities. Their nature and meaurement*. Mahway, NJ: Lawrence Erlbaum; Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory*. Cambridge: Cambridge University Press.
6. Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829–6833.
7. Thompson, P., Cannon, T., Narr, K., van Erp, T., Poutanen, V., Huttunen, M., et al. (2001). Genetic influences on brain structure. *Nature Neuroscience, 4*, 1253–1258; Chiang, M. C., Barysheva, M., Shattuck, D. W., Lee, A. D., Madsen, S. K., Avedissian, C., et al. (2009). Genetics of brain fiber architecture and intellectual performance. *Journal of Neuroscience, 29*(7), 2212–2224.
8. Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*(2), 171–191.
9. Devlin, B., Daniels, M., & Roeder, K. (1997). The heritability of IQ. *Nature, 388*(6641), 468–471; Richardson, K., & Norgate, S. (2005). A critical analysis of IQ studies of adopted children. *British Journal of Educational Psychology, 75*(3), 339–350.
10. This study involves groups of three – a pair of monozygous male twins (nearly 100% identical genes) and a brother (~50% gene overlap with twin). The title suggests that the twins were much more alike with respect to fMRI studies during problem solving. It does not point out that twins generally have much similar learning environments than do paired siblings. Also, twins generally learn to speak a bit later, on average, and they also are described as developing interpersonal communication schemes that only they perceive. It would be

important, therefore, to study large numbers of monozygous twins in which the differences were large in an attempt to tease out a better understanding of environmental influences. Koten Jr., J. W., Wood, G., Hagoort, P., Goebel, R., Propping, P., Willmes, K., et al. (2009). Genetic contribution to variation in cognitive function: An fMRI study in twins. *Science, 323*, 1737–1740.

11. Toga, A., & Thompson, P. (2005). Genetics of brain structure and intelligence. *Annual Review of Neuroscience, 28*, 1–23.
12. Plomin, R., & Kosslyn, S. (2001). Genes, brain and cognition. *Nature Neuroscience, 4*, 1153–1154.
13. Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown, and Company.
14. Colvin, G. (2008). *Talent is overrated: What really separates world-class performers from everybody else*. New York: Penguin Group.
15. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*. New York: Ballantine Books.
16. http://en.wikipedia.org/wiki/Matthew_effect (Accessed March 23, 2009); Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.
17. Wattie, N., Baker, J., Cobley, S., & Montelpare, W. J. (2007). A historical examination of relative age effects in Canadian hockey players. *International Journal of Sport Psychology, 38*(2), 178.
18. Another book related to talent acquisition that appeared just as this book was going to press is Coyle, D. (2009). *The talent code. Greatness isn't born. It's grown. Here's how.* New York: Random House. The title of this book as well as its content is consistent with the notion that innate talent is much more myth than fact.
19. Bayliss, D. M., Christopher, J., Baddeley, A. D., Gunn, D. M., & Leigh, E. (2005). Mapping the developmental constraints on working memory span performance. *Developmental Psychology, 41*, 579–597; Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, *89*, 63–100.
20. See Newcombe, N. S. (2002). The nativist-empiricist controversy in the context of recent research on spatial and quantitative development. *Psychological Science. 13*, 395–401; Kagan, J. (2008). In defense of qualitative changes in development. *Child Development, 79*, 606–1624.
21. Geary, D. C., & Huffman, K. J. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psychological Bulletin, 128*, 667–698.
22. Case, R. (1995). Capacity-based explanations of working memory growth: A brief history and reevaluation. In F. E. Weinert & W. Schneider (Eds), *Memory performance and competencies: Issues in growth and development* (pp. 23–44). Hillsdale, NJ: Lawrence Erlbaum Associates.
23. Kandel., E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (4th Ed.). New York: McGraw-Hill.
24. See Piaget, J. (2003). PART I: Cognitive development in children: Piaget: Development and learning. *Journal of Research in Science Teaching, 40*(Suppl), S8–S18; Vosniadou, S., Baltas, A., & Xenia, V. (Eds.). (2007). *Reframing the conceptual change approach in learning and instruction*. New York: Elsevier Science.
25. Vosniadou, S., Baltas, A., & Xenia, V. (Eds.). (2007). *Reframing the conceptual change approach in learning and instruction*. New York: Elsevier Science
26. Wollman, W. (1984). Models and procedures: Teaching for transfer of pendulum knowledge. *Journal of Research in Science Teaching*, *21*(4), 399–415; Wollman, W., & Lawson, A. E. (1978). The influence of instruction on proportional reasoning in seventh graders. *Journal of Research in Science Teaching*, *15*(3), 227–232.
27. Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
28. Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th Ed.). Englewood Cliffs, NJ: Prentice Hall; Moshman, D. (2005). *Adolescent psychological develo-ment: Rationality, morality, and identity* (2nd Ed.). Mahwah, NJ: Erlbaum.

29. Vygotsky's terminology includes the "zone of proximal development," or ZPD. (Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.).
30. Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
31. http://projects.coe.uga.edu/epltt/index.php?title=Piaget%27s_Constructivism (Accessed April 5, 2009).
32. http://projects.coe.uga.edu/epltt/index.php?title=Vygotsky%27s_constructivism (Accessed April 5, 2009).
33. For example see Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation.* (Vol. VIII, pp. 47–90). New York: Academic Press.
34. Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*(6), 1038–1063.
35. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.
36. Opdenacker, C., Fierens, H., van Brabant, H., Sevenants, J., Spruyt, J., Slootmaekers, P. J., et al. (1990). Academic performance in solving chemistry problems related to student working memory capacity. *International Journal of Science Education, 12*(2), 177–185; Johnstone, A. H., Hogg, W. R., & Ziane, M. (1993). A working memory model applied to physics problem solving. *International Journal of Science Education, 15*(6), 663–672.
37. van Merrienboer, J., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology*, *20*(3), 343–352.
38. Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer. See pp. 116–118.
39. Harp, S. F., & Mayer, R. E. (1997). The role of internet in learning from scientific text and illustrations: on the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89*(1), 92–102; Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187–198.
40. Cooper, G. (1998). Research into cognitive load theory and instructional design at unsw. http://paedpsych.jk.uni-linz.ac.at:4711/LEHRTEXTE/Cooper98.html (Accessed September 4, 2008).
41. Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.
42. The first issue of volume 21 of *Educational Psychology Review* included seven papers. The introductory paper provided an overview (Ayres, P., & Paas, F. (2009). Interdisciplinary perspectives inspiring a new generation of cognitive load research. *Educational Psychology Review, 21*(1), 1–9). Three papers within were of special note. Sweller has begun attempts to apply CLT to the issue of creativity (Sweller, J. (2009). Cognitive bases of human creativity. *Educational Psychology Review, 21*(1), 11–19). Renkl addresses worked examples and heuristic problems (Renkl, A., Hilbert, T., & Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychology Review, 21*(1), 67–78). Kirschner addresses issues connected to CLT and collaborative learning (Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks *Educational Psychology Review, 21*(1), 31–42).
43. http://en.wikipedia.org/wiki/Archimedes (Accessed March 23, 2009).
44. Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research, 57*, 51–67; Vosniadou, S., Baltas, A., & Xenia, V. (Eds.). (2007). *Reframing the conceptual change approach in learning and instruction*. New York: Elsevier Science.
45. Early in the development of the ULM we realized that "Ah, ha" moments are unusual, and that ordinary learning and ordinary conceptual change happen without drama. From Ohlsson: "It follows that nothing extraordinary happens at the moment of conversion. . .." Ohlsson, S.

(2009). Resubsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist, 44*(1), 20–40 (p. 32).
46. Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 1–33). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
47. Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., et al. (2004). Neural activity when people solve verbal problems with insight, *PLoS Biol* (Vol. 2). http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0020097
48. Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience, 30*, 57–78. As you read this review, think about attention as the premier slot in working memory and the rest of the slots devoted to deciding which among the potential slots is to be the focus of attention.
49. Robert Silberman (personal communication, August, 1984).
50. Thioux, M., Stark, D. E., Klaiman, C., & Schultz, R. T. (2006). The day of the week when you were born in 700 ms: Calendar computation in an autistic savant. *Journal of Experimental Psychology: Human perception and performance*, *32*(5), 1155–1168.
51. Ross, A. (July 24, 2006). The storm of style: Listening to the complete Mozart. *The New Yorker*. http://www.newyorker.com/archive/2006/07/24/060724 crat_ atlarge (Accessed March 23, 2009).
52. http://en.wikipedia.org/wiki/Expert (Accessed March 18, 2009).
53. Parker, E., Cahill, L., & McGaugh, J. (2006). A case of unusual autobiographical remembering. *Neurocase, 12*(1), 35–49.
54. http://en.wikipedia.org/wiki/Eidetic_memory (Accessed March 23, 2009).
55. Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
56. Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist*, *41*(4), 207–225.
57. Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
58. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*. (p. 5) New York: Ballantine Books.
59. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*. (p. 533) New York: Ballantine Books.
60. Gardner, H. (1993). *Multiple intelligences: The theory in practice. A reader*. (p. 51; p. 36) New York: Basic Books.
61. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*. (p. 537) New York: Ballantine Books; Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406; Howe, M. J. A. (1999). Prodigies and creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 431–448). Cambridge: Cambridge University Press.
62. Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown, and Company.
63. Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology, 87*(2), 287–309.
64. Gardner, H. (1998). *Extraordinary minds*. (pp. 149–152) New York: BasicBooks.
65. Cassidy, S. (2004). Learning styles: An overview of theories, models, and measures. *Educational Psychology, 24*(4), 419–444.
66. Cassidy, S., & Eachus, P. (2000). Learning style, academic belief systems, self-report student proficiency and academic achievement in higher education. *Educational Psychology, 20*(3), 307–322.
67. Krätzig, G. P., & Arbuthnott, K. D. (2006). Perceptual learning style and learning proficiency: A test of the hypothesis. *Journal of Educational Psychology, 98*(1), 238–246.
68. Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). Learning styles and pedagogy in post-16 learning: a systematic and critical review Available from http://www.voced.edu.au/td/tnc_79.71 (Accessed March 23, 2009).

69. We argue with the notion that our cellular makeup predestines us to a style of learning, or even makes on easier than another. There are specialized cells; as we were finishing our writing, still another was discovered: Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., & Moser, E. I. (2008). Representation of geometric borders in the entorhinal cortex. *Science, 322*(5909), 1865–1868. Most of the specialized cells feed into tissues that go on to store "memories" in the cortex. Not only is there compelling evidence for such cell types, but they most often are found when being specifically sought after their existence was predicted by a model – as in this recent report. There is a general storage geography for the brain. For example, visual content is stored at the back. Information from the left field of vision of each eye is stored on the right side. When one presses either the psychology or the biology, data supporting styles based upon genetic differences do not yet reveal themselves. Quite the opposite. Learning leads to development of "talents," and these talents appear to be reflected biologically in changes measureable with fMRI, etc. If there are style differences, they are small and most often readily overcome. As with any other difference, any inherent difference might be subject to a multiplier effect.
70. http://en.wikipedia.org/wiki/Homunculus (Accessed March 23, 2009).
71. Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. VIII, pp. 47–90). New York: Academic Press.
72. Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.
73. http://en.wikipedia.org/wiki/Homunculus_argument (Accessed March 23, 2009).
74. Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience, 139*, 105–118.
75. Hyde, J., Lindberg, S., Linn, M., Ellis, A., & Williams, C. (2008). Diversity: Gender similarities characterize math performance. *Science, 321*(5888), 494.
76. With respect to gender differences, "the past 5–10 years have witnessed a surge of findings from animals and humans concerning sex influences on many areas of brain and behavior, including emotion, memory, vision, hearing, processing faces, pain perception, navigation, neurotransmitter levels, stress hormone action on the brain and disease states." (Cahll, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience, 7*, 477–484.) Many of the gender differences reported are based in lateralizations; that is, something may tend to come in through structures on the left for men and right for women (or vice versa). One of the most obvious lateralizations one can observe is handedness. Some left-handed students require special equipment, such as student desks with writing tables on the left rather than right. As far as we know, there are no clear instructional strategies that differ based upon handedness. Does handedness matter? Well, think about this for a moment. About one in ten persons in the United States is left-handed. Of the last fourteen presidents of the United States, five have been left-handed and two others either left-handed or ambidextrous (http://en.wikipedia.org/wiki/List_of_left-handed_Presidents_of_the_ United_States (Accessed March 23, 2009)). The point is that, as with handedness, there may be something really important about instruction and gender that remains unknown to us. In *Brain Rules*, Medina devotes a chapter to gender differences. As we noted, many differences are related to lateralizations. Since no mention is made of handedness, and much of the writing is non-specific, we suspect that Medina is as unable as we are to nail down specific gender differences that might apply to learning. (Medina, J. (2008). *Brain rules: 12 principles for surviving and thriving at work, home, and school*. Seattle: Pear Press.).
77. Zeldin, A., Britner, S., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 48*(9), 1036–1058.
78. Geary, D., & Huffman, K. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psychological Bulletin, 128*(5), 667–698.
79. Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

80. http://en.wikipedia.org/wiki/Soar_(cognitive_architecture) (Accessed March 25, 2009).
81. Bereiter, C. (1990). Aspects of an educational learning theory. *Review of Educational Research, 60*(4), 603–624.
82. Ericsson, K., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245.
83. Schraw, G., Brooks, D. W., & Crippen, K. J. (2005). Improving chemistry instruction using an interactive, compensatory model of learning. *Journal of Chemical Education, 82*(4), 637–640.
84. Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. VIII, pp. 47–90). New York: Academic Press.
85. Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.
86. Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199; Simon, H. A. (1994). The bottleneck of attention: Connecting thought with motivation. In W. D. Spaulding (Ed.), *Integrative views of motivation, cognition, and emotion* (Vol. 41, pp. 1–21). Lincoln, NE: University of Nebraska Press.

Part II
Applying the Unified Learning Model

Chapter 7
Classroom Applications Overview

Because learning always involves the same big three factors (working memory, knowledge, and motivation – capacity, prior use, and current use), this book can serve as the intellectual basis for developing instructional programs in a research-oriented postdoctoral program at a major university, or a law college, or a pre-school. Because it deals with motivation so explicitly, the book implies opportunities for improvement even in situations already known for highly efficient training such as the military and some industries. Nevertheless, many teachers are probably saying to themselves, "Okay, so how does this fit in ***my*** classroom?" No two classrooms are exactly alike, and the range in classrooms is enormous. A fifth grade class in an inner-city, lower socio-economic status neighborhood where 50% of the children turn over during a school year, bears little resemblance to a fifth grade classroom in an upper-middle class socio-economic environment where only one or two children leave and one or two others come into the class during the year. The notion that there is such a thing as a context-free fifth grade in the United States is ludicrous. With this in mind, our intent is not to try to provide a comprehensive cataloging of specific instructional methods or teaching approaches for specific grades or settings. The power of the ULM is its parsimony. In the ULM, all effective classroom practice is anchored in the three principles of learning:

1. Learning is a product of working memory allocation.
2. Working memory's capacity for allocation is affected by prior knowledge.
3. Working memory allocation is directed by motivation.

These three principles provide a straightforward set of guides for any decision making about teaching methods at any age or grade level in any educational setting, from pre-school to graduate school, to business and industry, and to informal learning in places like museums.

For learning to occur, a student's working memory must be allocated to the learning task. Effective teaching must facilitate and attempt to optimize a student's ability to allocate her working memory capacity to the things we want her to learn. Students' prior knowledge can expand their working memory capacities and direct their working memory allocation. Effective teaching engages a student's existing knowledge and utilizes it productively. Teachers also must attend to possible

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_7,

negative impacts of prior knowledge, such as misperceptions that are discordant with what they are trying to teach.[1] Finally, students' working memory is allocated to learning to the extent that students are motivated to learn. Effective teaching must support and enhance students' motivation for learning. There are many ways that teaching can effectively enact the three principles of learning. The ULM says simply that teaching that follows the three principles of learning will be effective; teaching that does not follow these principles is very unlikely to be effective.

Beginning science students focus their attention on surface features of a problem, like whether it has pulleys or inclined planes, rather than on fundamental principles like whether it involves Newton's second law.[2] The ULM provides fundamental principles of learning in the same way as do principles of physics or chemistry or genetics. Instructional effectiveness comes from fidelity to the three principles, not from surface differences in media or methods. From the perspective of the ULM, we should not be debating, say, whether lecture or hands-on learning is a better instructional strategy. Instead, *we should be determining how well either approach focuses working memory allocation, engages prior knowledge, and provides motivation*. Either lecture or a hands-on activity could do a good or a bad job of these. Any other differences between them are simply irrelevant surface features.

The five rules of learning provide specific guidance for enacting the three principles. These say that effective instruction must do the following:

1. Direct student attention to the desired knowledge to be learned. Help students focus attention on relevant materials and avoid distractions through the learning environment, instructional materials, and connection to students' prior knowledge.
2. Provide necessary repetition. Provide multiple exposures to the knowledge to be learned and opportunities for recall and practice.
3. Facilitate connections. Provide ways for students to connect what they are learning to what they have previously learned in the class, what they have learned in other classes, and their other prior knowledge. Help them to construct meaningful connections between what they know and what they are learning.
4. Provide a learning environment that facilitates motivation. Recognize that learning can be difficult and provide support for maintaining students' effort.
5. Remember that learning is learning. Directing attention, providing repetition, facilitating connections, and providing motivation are the parts of good instruction. There are no short cuts; good teaching does not follow fads.

In the chapters that follow, we attempt to show how the ULM can be applied in real classroom settings.

Notes

1. Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 319–344). Mahwah, NJ: L. Erlbaum Associates, Inc.
2. Chi, M. T. H., Glaser, R., & Farr, M. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Chapter 8
Supporting Motivation

In Chapter 5, we detailed the underlying mechanisms of how motivation works to direct working memory allocation. We examined a set of core cognitive motivational constructs including goals, expectancies, and self-efficacy, along with interest and emotion. The discussion in Chapter 5 was directed at describing how these motivational constructs were represented cognitively and how they motivate attention and other working memory processing. We particularly noted that goals, expectancies, self-efficacy and to some extent interest and emotion were knowledge representations in long-term memory. They are primarily records of our episodic experiences in pursuing and achieving goals along with any associated emotional feelings, but also can be acquired and altered through vicarious experiences observing others and most importantly for this chapter, through teaching, feedback, and classroom practices. Like other aspects of applying the ULM to teaching, we cannot directly place knowledge, in this case the knowledge of goals, expectancies, and self-efficacy, personal interest, and emotions, into a student's brain. Teachers, however, can provide experiences that lead to students acquiring positive episodic motivational memories. Teachers also can structure classrooms in ways that help students set productive goals for learning, provide students with positive emotional experiences, and utilize interest to direct students' attention. Teachers can use feedback to help students develop productive means-ends expectancies and build self-efficacy for their academic achievement.

We can't provide a comprehensive review of all the effective classroom practices that have been identified for building motivation. We instead will focus on classroom practices for building the key motivators that we discussed in Chapter 5, as well as discuss how teachers can overtly provide motivational direction to students. It should come as no surprise that "teaching" connected with those things we label as motivation will involve attention, repetition, and connections. Although some episodic learning will be easy, semantic learning with respect to motivation will require effort. In the end, what is learned about motivation will be learned through the same neural mechanisms as would knowledge more often thought of as "learned in schools", such as an understanding of democracy or approaching solutions to differential equations.

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_8,

Learning Goals

We noted earlier that goals were central to all purposeful, "controlled" processing. Because learning in educational settings is purposeful, learning in schools and other educational settings is goal directed. Students are always pursuing some goal in school. For an educational experience in school to be successful, however, students must be pursuing goals that are concordant with the learning outcomes that teachers want. *Goals motivate their own attainment*. Once we have a goal, we exert cognitive effort to attain that goal. Goals form the foundation for student motivation in school and for where their working memory capacity is being allocated. Goals matter. When they interviewed young students about their assignments, Anderson and her colleagues found that those who didn't have a clear idea about their goals for learning tasks responded, "I don't know what it means, but I did it."[1] Their goal was to complete the assignment, but not to learn anything while they did it. Odds are that not much learning took place.

Work in goal theory has identified clusters of *goal orientations* in the academic setting that underlie most achievement motivation. The initial distinction, which emerged from work by both John Nichols and Carol Dweck, was made between what are called learning or mastery goals and performance goals.[2] Learning goals are goals directed at learning new knowledge or mastering a task or problem. Performance goals are goals directed at demonstrating ability or doing especially well in relation to others. These clusters subsume a rather broad array of specific goals that students might have for specific assignments or tasks. The idea behind these clusters is that they constitute relatively stable *orientations* or general tendencies to set certain types of goals. From the perspective of the hierarchical structure of motivation, these orientations are general academic goals that would be more stable than the specific goals being set for individual assignments or activities. While it was originally thought that these orientations were diametrically opposed (that is, one had either a performance or a mastery tendency), research has shown that they are essentially independent.[3] Students may be pursuing one or the other or both. Subsequent work has also identified *work avoidance* goals, which are goals to get through the class or task with as little time and effort as possible,[4] and their opposite *task goals*, which are goals to complete the task well, such as getting a high grade but without any expectation of learning or comparative performance to others.[5]

From the perspective of the ULM, the most productive goal for a student to have in an educational setting is a learning goal. This is a goal that will direct working memory allocation to learning as opposed to some other outcome.

Specific learning goals result from a learning goal orientation. The foundation for effective teaching, instruction, and student motivation is establishment of conditions that support student development of a learning goal orientation and students setting specific learning goals for their classroom engagement and studying.[6] Students may be pursuing other goals along with a learning goal, but the learning from pursuit of a performance or task goal will be "incidental" to the goal and may or may not occur or be lasting.[7] The most likely way for meaningful learning to happen is when a learning goal is being purposely pursued.

This was made clear to one of us many years ago when he was teaching an introductory course in an industry education program. The course was part of a certification sequence. The company offered a $50.00 bonus (when that was worth a lot more than today) for employees who completed the sequence. He asked students in the course why they were taking it. There were three answers: (1) to get the $50.00 bonus, (2) to learn knowledge that would help them do their job better, and (3) to just learn as much as they could about the industry. On the final exam, those who were in the course for the $50.00 didn't pass, those who were in it to help with their job passed just above the 70% criterion, and those who just wanted to learn as much as possible all passed with scores above 90%. This clearly unscientific study was done long before the notion of learning goal orientations was formalized in the literature, but it seems clear that those who had learning as a goal learned more. Interestingly, those who had a focused practical goal of gaining job related skills did not do as well as those who just wanted to learn as much as they could, in general.

Learning goals are *not* the same thing as "intrinsic motivation" or interest. *A learning goal does not require that you are interested in or enjoy what you are learning*. One of the authors is a statistician who really doesn't like math. He had, and still has, no desire to study statistics for the sake of studying statistics. He does like doing research, however, and he set goals to learn statistics so that he could properly analyze his own research data. Much of what we want students to learn in school is like this. It is valuable because it is connected to other desired things. Too often learning goals are confused with learning for the sake of learning. We spend considerable time trying to convince students that they should like math or science or reading or history. We do this because we expect that, if we get students to like the subject, they will set goals to learn it. It is true that students who are intrinsically motivated by their personal interest in a subject will likely set learning goals. But it is neither certain that we can get all or even most students to like or develop personal interest in some subjects, nor certain that liking a subject in and of itself will automatically lead to a learning goal. Enjoying a subject can just as easily lead to a task goal of doing an activity for the fun of it. Consider a student in a hands-on science project who expresses an enjoyment for working with his hands to build a research apparatus, but doesn't really care about learning from the results of the experiment that the apparatus is for.

At the core of successfully motivating students is getting them to set a learning goal. Whatever goals your students may have, you can set learning goals for your students in your classroom and achieve some "buy in" for those goals. Pressley and his colleagues describe primary grade classrooms where teachers established that learning was a goal, and that children "bought into" that goal.[8] There is no certain classroom approach for insuring that students set learning goals, but you can bring learning goals to the forefront of their thinking as you start and end a lesson or a unit. One way to engage students in the topic to be studied is to use a KWLQ chart ["When I say the word 'insects', what do you *Know* about them?" "Now tell me what you *Want* to learn about 'insects'." "What new information did you *Learn* about insects now that we read about them and watched the video?" "What unanswered or new *Questions* do you now have about insects? Where or how might you find the answers to those questions?"][9] It can be a way to introduce the topic and guide your initial discussion, while at the same time having students actively generate learning goals for the lesson and beyond it.

While competition and normative, curve grading can cause students to focus on performance goals of doing better compared to other students, they do not preclude learning goals. Highly motivated, self-regulated students generally pursue both learning and performance goals. Competitive classrooms and normative grading, however, can adversely affect students who are doing poorly. They can lead to students focusing on what are called performance avoidance goals of trying not to look dumb and work avoidance goals that produce negative emotions that interfere with learning goals.[10]

Bereiter and Sweller both have noted that students often approach schoolwork with a task goal.[11] They usually want to do well, but assignments, projects, homework, and so on are regarded as tasks to complete rather than opportunities to learn. Bereiter has suggested that the key to getting students to pursue learning goals rather than task goals is to allow them freedom to construct their own knowledge and meaning.[12] This "knowledge building" approach is one of the five motivational profiles we will describe in Chapter 12. It has also been central to approaches based on collaborative learning communities[13] and computer supported collaborative communities (CSCL).[14] Some research has shown that establishing these types of collaborative communities does increase students' use of a knowledge building approach. However, just providing cooperative learning groups may or may not be effective. In one school that had mandated that all instruction be in cooperative groups, students began to see group work as just a task to complete rather than as a support for learning.[15]

Sweller has taken a somewhat different approach focused on aspects of specific learning tasks. He notes that, when students are given a problem to solve, this often creates a task goal of getting the problem completed correctly. This causes students to direct working memory allocation to finding problem solutions, often through trial and error, rather than focusing on learning the underlying principles that the problem was supposed to illustrate. He found that students actually were slower to learn more sophisticated problem solving skills when they had developed successful rudimentary strategies for getting problems solved. In response, he used worked examples to direct students' goals toward learning the underlying problem-solving algorithms rather than toward task completion.[16] Worked examples are a form of observational or vicarious learning. By using worked problems instead of practice problems to show students how to do something, they can set goals to learn without concern about their performance or getting the task done correctly.

These approaches to supporting learning goals are based on keeping students focused on learning or developing their own mastery rather than on competing normative performance or task goals. From our previous discussion of goals on motivation (Chapter 5), it is clear that keeping students focused on a learning goal requires that students succeed in learning in order to develop an expectancy of learning success. This suggests that what have been called mastery learning approaches are likely to be especially useful for fostering learning goals. By establishing a clear learning goal and allowing students opportunity, including multiple chances, to succeed, mastery approaches keep the student focused on continued learning and mastery of material.

Teachers can support the idea that success is possible through verbally assuring students that they can succeed in learning. Classes where teachers do this have more students who set learning goals.[17] As an added bonus, as teachers create learning environments that foster success and employ specific success messages, students are more likely to perceive positive teacher-student relationships – a situation that predicts positive student emotions toward school.[18]

While it would be optimal to have students setting their own learning goals, all too often teachers have students who in a least some instances "don't care" whether they learn the material; they have a *learning avoidance goal orientation*. This is especially prevalent in "required" courses in a structured curriculum; there are those synthetic organic chemists who wonder out loud why they must pass a course in chemical physics. We also see this in some students, especially those in some ethnic or religious groups who view school subjects as irrelevant and possibly even oppressive. These students may actively resist learning what is being taught. Unless the student is truly one of the apathetic students as described in Chapter 12, he knows that, even if he doesn't care about certain content, he must succeed at that content. This usually produces a task goal orientation where success is typically measured in terms of a letter grade reported on a report card or transcript rather than in personal knowledge growth.

Teachers can address this through one of their most significant tools for directing student goals: written or verbal feedback on their assessments and graded assignments. Regardless of any personal goals, students will almost always treat course objectives and graded work as goals to be pursued. Although extensive standardized testing and a focus on "standards" have given assessment a bad name in many education circles, assessments can play a positive role in focusing students' learning goals. The key is to create assessments and assignments that require students to learn at a more than just superficial level to be successful. If assessments can only be answered by having multiple repetitions and practice, developing chunks with extensive hierarchical connections, and dynamically constructing novel answers to problems, students will have to learn, even if they only want to "get through the course." In effect, the assessment can set the learning goal for the student.

Belief in Effort

As we discussed earlier, students are not likely to set goals unless they believe they can achieve them. While expectancy of success is the primary expectancy that affects pursuit of a goal, students often are encountering new material and subject matter. Therefore, they often have little direct experience on which to develop strong expectancy of success for their learning goals. Their success is an unknown. They typically draw on their past experience with similar subjects or school work in general to develop an initial expectancy of success. This is one key reason why a learning goal orientation is so critical. Dweck and Leggett showed that students who have learning goals pursue new learning tasks and shrug off failures as learning opportunities.[19] They are focused on the expectancy for gaining new knowledge. This orientation is highly productive in the typical school experience of continually encountering new material. Similar results have been found by others and with adults as well as children.[20] If a student puts forth effort studying, he can anticipate a relatively high expectancy of success for learning something. In contrast, success at doing better than others or getting a high grade is considerably less certain.

As embodied in the fourth rule of learning, pursuing these learning goals is difficult. It is going to take effort. At its core, motivation is about effort. This is true in both the immediate context of focusing working memory attention on a specific learning task as well as in the broader context of developing expertise in a chosen discipline, something that can take many years of hard work. While many strategies may work in a given context, all will require that the learners expend effort. Deliberate "controlled" or self-regulated learning requires effort. We have to be willing to focus our concentration on what we are learning in the face of distraction, to spend the time to review and go over the material multiple times, to practice skills repeatedly, to make the connections necessary to create chunks, and so on. Most of the time, these require us to stick to it and persevere more than anything else.

Putting forth effort requires that we have a means-ends expectancy that effort will achieve our learning goal.[21] We have discussed the hierarchical nature of motivation. Within means-ends expectancies, a small set of four general means-ends expectancy clusters have been identified. These have also been called causal attributions in the literature. The four are effort, ability, task difficulty, and luck. Considerable research has shown that attribution to these produce differential motivational and emotional outcomes.[22] Critical for our discussion here, putting forth effort requires that students see effort as a viable means for achieving their learning goals. If students believe that luck or the difficulty or ease of the assignment determine how much they learn, they may not be motivated to put forth as much effort as they need to succeed.

Effort interacts strongly with how students perceive ability. Dweck and Leggett found that children who believe that their ability or intelligence is fluid or changeable through learning, what they call an *incremental view of intelligence*, set learning goals and increase effort in the face of difficulty. Children who believe that their ability is fixed, like classic definitions of IQ called an *entity view of intelligence*, set performance approach or avoid goals depending on their success or failure. They decrease effort in the face of difficulty. Dweck and her colleagues have further confirmed that incremental views of intelligence foster means-ends beliefs in effort that are associated with setting learning goals.[23]

Consideration of how students and teachers perceive ability is important because studies have shown that children, parents, and teachers in the U.S. predominantly hold entity views of intelligence.[24] This results primarily from an emphasis in U.S. schools and the broader society on things like IQ, giftedness, and multiple intelligences. When people talk about athletes like Lance Armstrong or Michael Phelps or Tiger Woods, they often attribute their successes to their “talent.” The years of effortful practice that these athletes have engaged in are ignored. When we look at those who are “the best” at what they do, be they athletes or engineers or scientists or musicians, what we find is that the top performers practice and work most diligently. Benjamin Bloom, best known for the *Taxonomy of Education Objectives*, made a similar observation in a study of talented people described as the top 120 Americans in six different professions.[25] He concluded that almost any person could accomplish what these successful people had if the conditions were right, reiterating the role of effort and learning in achievement rather than special native ability. In *Outliers*, Gladwell stresses that the most successful among us have invested 10,000 hours of intense study if not more.[26]

Research has shown that when students are told that talent or genius are due to “natural” abilities, they develop entity theories of intelligence. Entity views essentially communicate to children that “you are either good at it or you’re not” or “you’re born with the talent” and that nothing you can do can change it. This fosters in U.S. students a means-ends belief that ability rather than effort is the primary cause of success. How often do we hear students say something like, “Well I’m just not good at math [or science or reading] so there is no sense trying too hard”? These entity views can be established and reinforced by teachers praising students for doing well because they are smart. Worse yet was when one of the authors heard a parent state in front of his child at parent-teacher conferences, “Well, I was never good at math, so she must have inherited that from me. I’ve done fine and so will she.” Entity views impact not only students’ willingness to set learning goals and devote effort to them but also whether students take more demanding courses like advanced math or science. It also causes students to focus more on performance goals or task goals, where they “demonstrate” how smart they are by doing better than others or getting high grades, rather than on goals of learning to further develop their knowledge and skills. Different world views may result in different implicit beliefs about intelligence. This view contrasts sharply with those of students, parents, and teachers in Asian societies like Japan and China where an incremental view is predominant. There students are told that regardless of what their current ability is, they can get better through hard work and that success depends on effort.[27]

As discussed in Chapter 6, the ULM reframes crystallized intelligence as knowledge. Furthermore, in the ULM, that knowledge is always changeable and increasable through learning. *The ULM itself is an incremental model of intelligence and ability*. This suggests that teachers who adopt the ULM and talk to their students about learning, effort, and ability in ULM terms will likely foster an incremental belief in their students. Recent studies have confirmed that belief in the incremental view can be changed through instruction. Students, many of whom were “at-risk”, who were taught or did computer-based lessons about how the brain is changed by learning, had more incremental views of intelligence and ability and subsequently greater growth in their school achievement than control groups taught only good learning strategies.[28]

As we discussed in Chapter 5 on motivation, means-ends expectancies reflect one's experience with achieving or not achieving the goal with those means. Students will come to have a means-ends expectancy for effort to the extent that effort has, in fact, achieved their goals. If students are putting forth effort and not succeeding in learning and completing assignments successfully, they will begin to see these as resulting from things other than effort. Students with a history of failure can see both learning and grades as primarily dependent on luck or whether they have easy assignments to compensate for their lack of ability. As a result, they are not motivated to put forth effort. It is critical, therefore, that classroom assignments be achievable by students through their effort. Again, mastery learning approaches that structure classroom assignments in ways that foster achievement through effort support development of means-ends expectancies for effort. Teacher feedback that success is due to effort can further increase these effort beliefs. Teachers at any level need to convey to their students that they have the capability of being successful and that success is realized through effort.[29]

Goal Value and Outcome Expectancies

We have talked about the importance of setting learning goals, but school is about more than just learning and mastering the material being taught. We all likely hope that our students develop personal interest in the various content areas we teach, like math or science or literature or the arts. We have come to value mastery of these for their intrinsic value. The fact remains that, however, school is about many more outcomes. Students value learning goals in school not only for the sake of learning the material but also for the other goals and outcomes that this knowledge might have utility for obtaining.

The ULM is very unique in that it focuses on learning as a cognitive process in real time. Learning is seen as the result of how students are directing their attention moment by moment and how this attention is engaged repeatedly. Most of our focus in the chapter so far has been on how motivation works in the context of this active learning process. This is immediate and short term. Social cognitive theorists have called goals at this level proximal or short-term. When we previously talked about getting students to set learning goals, we were looking at short-term goals for learning from ongoing engagement with learning materials. We emphasized this because, as we have previously discussed, learning is the result of the way working memory is allocated during real time interaction with what is being learned.

Short-term goals for learning a specific assignment or doing a specific homework problem or listening to a specific lecture must ultimately be connected to more long-term desired goals and outcomes, so-called distal goals. As we discussed in Chapter 5, these connections between our immediate short-term learning goals and more long-term outcomes that these might lead to are *outcome expectancies*; the likelihood that achieving our learning goal will result in other outcomes that we desire. Ultimately, if students are going to be motivated toward school and learning, they need to see what they are doing in school as relevant for what they hope to achieve in life. While they may value immediate learning goals for gaining knowledge, they also derive value for short-term learning goals from the outcome expectancies for long-term life goals.

Students value goals that are directly relevant to their daily lives. Often teachers are told to stress to students how useful what they are learning will be or to connect lessons to students everyday life. Many students struggle to find relevance in what they are learning. This is especially true for students in disadvantaged home situations. They may be dealing with needs at the physiological and safety levels of Maslow's Hierarchy that we discussed earlier.[30] Even for students from advantaged backgrounds, there often is little immediate relevance for algebra, advanced science courses, or literature in their day-to-day living.

The outcome expectancies for what students are learning in school are most likely more long-term. The goals that school can lead to may be well into the future for students in the early grades and can be long-term for college undergraduate and even graduate students. Outcome expectancies for goals that are long-term are also called *perceived instrumentality*, the belief that current behavior is a prerequisite to future goal attainment. Research has shown that for at-risk college students in a remedial mathematics course, believing that the course was instrumental to achieving their other academic goals led to greater motivation and better academic performance.[31]

The extent to which students can connect their short-term learning goals to future outcomes is affected by their *future time perspective* (FTP); their personal, individual structured representation of the future.[32] As part of their future, students must see attaining future goals as valuable and viable; what Markus has called the "Future Possible Self."[33]

While teachers can't directly change a students' overall future time perspective, instructional research has shown that students can increase their outcome expectancies that their learning in school is instrumental for future goal attainment. Oyserman and her colleagues developed the *Pathways for Youth* project, a nine-week, after-school program for urban middle school students.[34] Adult guides worked with students to complete seven steps: (1) envisioning possible futures for themselves, (2) conceptualizing those futures as goals, (3) constructing a path for goal obtainment, (4) making explicit connections between present educational activities and the valued future goals, (5) discussing possible roadblocks and forks in the path, (6) brainstorming strategies for managing imagined future obstacles, and (7) interviewing successful adults from the community about their own strategies for reaching goals. Results showed that, relative to students in a comparable group, students who participated in the program "... reported more bonding to school, concern about doing well in school, 'balanced' possible selves, plausible strategies to attain these possible selves, better school attendance, and for boys, less trouble at school"[35]

Oyserman's intervention embodies three foundational strategies for getting students to develop outcome expectancies for important possible future goals. First, students need to know that the future goal exists, be it further education or a career.

Students may have only limited knowledge of what is possible. While there are some goals that can't be reasonable for most of us (winning a Nobel Prize in Economics, being elected President of the United States, or winning the US Open Golf Tournament), all are potential goals: *someone* will achieve the goal. There are other laudable goals that almost anyone could potentially achieve. For example, nearly anyone could become an acknowledged expert in economics or hold elected political office or be a golf teaching pro. But, to pursue any of these goals, students need to know they exist. They need to know that there are pastors and physicists and painters and pharmacists and pianists and professors and so on. The more possible careers and outcomes they know about, the greater the potential for developing outcome expectancies for them.

Second, students need to see these goals as something they could possibly achieve: a future possible self. Many students do not see the outcomes that school could produce, such as further post-secondary education and technical careers, as possibilities for them. This is especially prevalent in minority and immigrant populations and among those in the lower socio-economic groups. It is not just economics that interferes with seeing these outcomes as possible. It can be a life situation that is sometimes by necessity so focused on immediate needs that the students can't see beyond their current environment to envision themselves in a better future place. They may also lack role models for possible careers and educational paths. As in Oyserman's intervention, teachers need to help students see their possible future selves and translate possibilities into goals to be pursued. Recently, a two-year follow-up to a subtle intervention involving structured writing assignments with minority students aimed at reducing the psychological threat of being negatively stereotyped showed long-term positive effects such as significantly improved grade point averages.[36]

Third, students need to develop outcome expectancies for these possible future goals. They need to lay out overtly the instrumental path from where they are now to these goals, including intermediate steps for taking the classes and other activities necessary to prepare themselves for each successive step. The connection between "right now" and the long-term future outcome is tenuous, but can be made concrete by establishing each successive step as the next short-term learning goal.

Self-Efficacy

One can recognize the value of a learning goal, see the outcome expectancy and instrumental path that connects this short-term learning goal to a desired outcome, understand the means for achieving the goal, but still not be motivated to set and pursue that learning goal. The missing motivational piece is self-efficacy. *Students will not set and pursue a goal unless they feel confident that they can do what is needed to achieve it*. Self-efficacy impels us to act. If we have high self-efficacy, we take action. If we have low self-efficacy we don't, even if we do know what to do. If self-efficacy is absent, all of these other motivators are not going to result in action or goal pursuit. People don't pursue goals they don't think they can achieve.

This means that developing students' self-efficacy is a high priority for teachers. Self-efficacy is grounded in students' actual experience of success and failure. If students have tried to learn something and failed, they have lower self-efficacy for the next time. If they have had success in learning, they will have higher self-efficacy for the next time. Establishing success is critical for building student self-efficacy. This, like most of the other motivators we have discussed, is well served by mastery approaches that allow ample opportunities for students to succeed.

There are many excellent resources that already exist on how to build self-efficacy in classrooms.[37] We recommend consulting these. As a quick summary, beyond mastery learning experience, self-efficacy can be increased vicariously by seeing peers with whom the student identifies success in learning or by genuinely deserved praise and feedback to students that their knowledge and capabilities are improving.

The most important consideration in trying to boost students' self-efficacy through praise and feedback is the way students view ability and effort. As we have noted previously, in the U.S. most students have an entity view of ability and intelligence as built-in and unchanging. One consequence is that they see effort and ability as diametrically opposed. If you have ability, you don't have to work hard; if you have to work hard, you don't have ability. Therefore, when students receive feedback and praise for working hard or putting forth effort or that their success was due to effort, all things we have previously advocated as ways to increase a means-ends expectancy for effort, they can interpret this as meaning they lack ability and have decreases in their self-efficacy as a result.

One way to break this cycle is through the previously discussed focus on getting students to adopt an incremental view of intelligence and ability. Self-efficacy is increased by feedback that one has or is increasing in ability. This is productive only if students understand that ability is incremental and increased by learning. When students have a ULM view of learning, ability and effort are no longer opposed; they are symbiotic. Ability is increased by effort and more ability means less effort in the future or the ability to tackle something more challenging. Self-efficacy is increased by both ability and effort feedback. Feedback that focuses on the process of learning and improving, coupled with feedback that ability is incremental, boosts self-efficacy more than feedback that focuses only on performance or outcomes.[38]

A teacher might think that telling a student he/she has done good work is enough, but research shows us a more nuanced picture. When interviewed about their careers, male and female scientists respond differently. For men, simply indicating that they have achieved mastery seems sufficient to enhance self-efficacy. Women, in spite of being told they have done well, may still see themselves as not being especially good at what they have done. Zeldin et al. report:[39]

> Analysis of 10 narratives revealed that mastery experience was the primary source of the men's self-efficacy beliefs. These results are compared to those from Zeldin and Pajares' earlier study[40] involving women in STEM careers. For women, social persuasions and vicarious experiences were the primary sources of self-efficacy beliefs.

In elementary school writing, girls showed greater relative self-efficacy than boys. These gender differences emerging from research (whatever their source may be) suggest that teachers should address both performance and self-efficacy when commenting to students. Remember, self-efficacy is not something you can teach, but it is something you can influence. One of us was considered to be an effective recruiter to careers in chemistry. The "pitch," always made after a student had done something well, was the same. The first question was "What's your major?" No matter what the response was except for chemistry, the next statement was, "Well, always remember, you're good at this and if you ever think you might want to do it for a living, stop by and see me." At the time this was done, there was no sense that it had research support, and it always was done in a gender-neutral manner with the initiating event always based upon the student's performance.

The second consideration in managing self-efficacy is that students will invariably compare their achievement to that of others. Self-efficacy is derived from normative comparisons as well as mastery. Students who see themselves as clearly behind or not achieving as well as other students will sooner or later begin to doubt their skills and lose efficacy. Because one cannot become an expert at everything, differences in performance and achievement are inevitable. While Bloom noted almost everyone could potentially become a high achiever in any field, it is equally true that they cannot become high achievers in all fields. It probably isn't productive for 5' 2" college student to follow a dream of becoming a player in the National Basketball Association and spend his time and effort developing basketball skills, even though there is no reason that he could not become just as "skillful" as a 6'10" player. The odds are probably too great for the effort to pay off.

In P-12 education, the reality is that nearly all students, with the exception of those with *severe* cognitive or developmental disabilities, can master all of the subject matter being taught. All students can become competent in reading, writing, math, social studies, science, arts, and physical education. While someone will be "the best" at each of these, everyone can be successful. Students need to be helped and guided to base their self-efficacy, not on how they compare to others, but on how they are gaining competence. Too often teachers and parents assume that children's potential is limited, usually because they, themselves, have a view the ability is somehow fixed at birth or by status (an entity view). This can lead to lower expectations. A Member of the U. S. House of Representatives recalls being told before completing high school that he should not aspire too much in the way of a professional career, so he quit school. After service in the U. S. Army where he gained a more realistic idea of what he was capable of achieving, he returned to complete high school, college, and law school. One of us had a doctoral advisee who revealed in a moment of confidence that his father always told him he would fail, and that had cast a cloud on the student's life view. Today, that person is a very successful academician, one who always seems to find a way to encourage students.

While it is well that students can overcome negative expectations if they have enough personal self-efficacy, it would be better if we did not impose negative expectations in the first place. By following the principles of the ULM, teachers provide a much more empowering view of students' ultimate capabilities and how learning through serious effort can achieve high levels of competence and performance. Expressing these beliefs will build students' self-efficacy that they, in fact, possess the means to achieve their goals.

Interest

Both situational and personal interests are relevant to the classroom. Situational interest directs attention. Things that are novel generate interest. Several of us once were connected with a project in which a zoo sent certain animals from their collection into classrooms. This practice afforded classroom opportunities for teaching and learning, and also addressed the zoo's problem of housing the animals during months when the zoo exhibits were closed to the public. One unexpected discovery resulted from this program. Whatever animal was brought into the classroom, that animal could serve as a learning center. During writing, for example, all animals served as effective story starters. What was noticed, however, was that unusual animals (legless lizards, ferrets) seemed to catalyze "more scientific" writing than did ordinary animals (rabbits, doves). Novel animals generated more interest.[41]

Because situational interest is a property of the environment, in particular of the classroom and learning materials and activities done, situational interest is something that a teacher or instructional designer can manipulate directly. Interesting instructional materials draw attention. Allocating working memory through focused attention is a prerequisite for learning, and interest can play a significant role in directing students to attend to relevant instructional content. Generating interest is often cited as a reason for incorporating multimedia or hands-on activities into the classroom. These can be novel and generate situational interest.[42] Consider how teachers can generate situational interest as they open a lesson. As a former middle school teacher, one of the authors taught a lesson that included reading Sandra Cisnero's short story *Eleven* and analyzing the literary symbolism found in it.[43] To begin the lesson, she would silently set a grocery store bag in front of her class and take out a ragged red sweater, a deflated balloon, an onion, and a metal Band-aid box with a coin in it that she rattled, and asked the students to predict what these things might have to do with the short story they were about to read. Setting the stage for instruction like this created situational interest in the upcoming task, established a purpose for reading, and provided a platform for some lively focused discussion on the Cisneros' choices for symbolizing the main character's experiences – all important components for successful literacy learning.

There are problems with designing instruction to generate situational interest, however. First, virtually nothing has universal situated interest for everyone. While we can assume that some things are likely to be novel for almost all students, such as the unusual animals in our zoo example or chemistry or physics demonstrations of phenomena students haven't seen before, we can't assume that even these are necessarily interesting for everyone. The same caveat holds for assuming that computer video or animations are always interesting. This holds for activities as well. Assumptions that all students are interested in hands-on work or in "game-like" computer learning environments will certainly be wrong. Our caution here is not to say that teachers and instructional designers should avoid trying to embed interesting novelty in their classrooms or materials. Rather, we caution against assuming that these will always be "intrinsically" or automatically interesting to students.

It's also not universally clear how students will "take" something. During a visit to a regional school noted for its academic excellence, one of us who never wears a cap noticed a baseball cap at a teacher's desk saying, "Head Coach, Physics." Months later, during a professional development context, he suggested giving science teachers such hats. Many teachers, especially women, viewed this as "sexist." It wasn't that such a hat would be neutral; it was that it would be viewed by some as offensive.

Second, situational interest can generate what have been called seductive details. Situational interest directs attention. But what if the thing that is interesting isn't what we want students to learn. One of the authors was involved in a study of eighth-graders reading their history textbook. One of the passages had a picture of women suffrage marchers from the 1890s. When asked what they recalled from the passage, the majority of students mentioned the women in the picture. But, they didn't remember what they were marching for (the right to vote), they recalled that they looked funny or sometimes less flattering terms. The students focused on their unusual appearance rather than what they were supposed to represent.

Because situational interest is so powerful for directing attention, novelty and other ways of generating interest need to be carefully planned and thought out. If teachers or instructional designers are going to use interest as an attention director, they need to be sure that the situational interest draws attention to instructionally meaningful content. Because students remember what they attend to, we need to be sure that we are making them attend to what we want them to learn, and not to some distracter.

In contrast with situational interest, personal interest sustains long-term engagement with learning in personally meaningful subject matter domains. As such, we can think of personal interest as emerging from sustained pursuit of learning/mastery goals in a discipline. Because personal interest is unique to each student and is based in the student's own emotional and cognitive reactions to their learning and developing expertise, there is little that a teacher or instructional designer can do to directly influence it. But, because it emerges out of growing competence and growth in a discipline, teachers can foster personal interest by following our previous recommendations for focusing on learning goals, belief in an incremental view of intelligence, belief in effort, and development of self-efficacy. As students build proficiency in a subject, they will begin to develop personal interest in knowing more about that subject.

In addition to building proficiency in a subject, teachers can harness personal interest as a means of providing opportunities for skills practice. For example, the best way to build reading fluency and enhance vocabulary development is for children to read a lot. This gives them multiple exposures to high frequency words, leading to automatic word recognition: the "glue of our language." Increased reading also apprentices them into the language of a particular topic and enhances their understanding of more precise and more generalized word usage. One way to entice students into increasing the amount of reading they do is to have them read self-selected books and websites on topics that are of personal interest to them. Thus, taking advantage of growing personal interests in this way can result in multiple facets of literacy learning.

Discouraged Terms

The comedian George Carlin was famous (notorious) for a list of seven words one couldn't say on television. Indeed, this list made it to a U. S. Supreme Court case.[44] There probably should be a list of terms that are banned in your classroom. For starters, how about:

- Talent – discouraged because it implies something you are born with rather than something you develop through learning and effort.
- Smart – discouraged because it tends to overemphasize the role of ability in learning and implies that effort cannot overcome ability deficiencies when, in many, if not most cases, it can.
- IQ – discouraged for the same reason as smart is discouraged.

Summary Thoughts on Motivation

Motivation directs working memory allocation. The goal of motivation is to direct students to focus working memory on the learning task. In this chapter, we have outlined some basic ways teachers can facilitate motivation. These are:

1. Focus students on setting learning goals, rather than performance or task goals.
2. Help students develop an incremental theory of intelligence and ability by basically teaching them about the ULM.
3. Foster students' belief in effort through feedback and mastery experiences.
4. Help students build outcome expectancies and specific instrumental paths that connect what they are learning to future goals and outcomes.
5. Build students' self-efficacy through mastery experiences, feedback focused on the process of learning, and encouragement that effort will build competence.
6. Utilize interest and novelty in ways that support allocation of attention to relevant learning materials rather than distracters.

Following these suggestions will help students maintain motivation and build positive self-beliefs.

Notes

1. Anderson, L. M., Brubaker, N., Alleman-Brooks, J., & Duffy, G. C. (1985). A qualitative study of seatwork in first-grade classrooms. *Elementary School Journal, 86,* 123–140. (Quote from p. 132.).
2. Nicholls, J. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91*, 328–346; Dweck, C. S., & Leggett, E. L. (1998). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.

3. Fryer, J. W., & Elliot, A. J. (2008). Self-regulation of achievement goal pursuit. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 53–75). New York: Erlbaum/Taylor & Francis Group.
4. Meece, J., & Holt, K. (1993). A pattern analysis of students' achievement goals. *Journal of Educational Psychology, 85*, 582–590.
5. Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology, 85*, 541–553; Ng, E., & Bereiter, C. (1991). Three levels of goal orientation in learning. *Journal of the Learning Sciences, 1*, 243–271.
6. Dweck, C. S., & Master, A. (2008). Self-theories motivate self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 31–52). New York: Erlbaum/Taylor & Francis Group; Meece, J., Anderman, E., & Anderman, L. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, *57*, 487–503.
7. Ng, E., & Bereiter, C. (1991). Three levels of goal orientation in learning. *Journal of the Learning Sciences, 1*, 243–271; Linnenbrink, E. A. (2005). The dilemma of performance-approach goals: The use of multiple goal contexts to promote students' motivation and learning. *Journal of Educational Psychology, 97*(2), 197–213.
8. Pressley, M., Kersey, S. E. D., Bogaert, L. R., Mohan, L., Roehrig, A. D., & Warzon, K. B. (2003). *Motivating primary-grade students*. New York: Guilford Press.
9. Schmidt, P. (1999). KWLQ: Inquiry and literacy learning in science. *Reading Teacher, 52*(7), 789–792.
10. Shell, D. F., & Husman, J. (2008). Control, motivation, affect, and strategic self-regulation in the college classroom: A multidimensional phenomenon. *Journal of Educational Psychology, 100,* 443–459.
11. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285; Bereiter, C., & Scardamalia, M. (1989). International learning as a goal of instruction. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 361–392). Hillsdale, NJ: Erlbaum.
12. Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of: The learning sciences.* (pp. 97–115). New York, NY, US: Cambridge University Press.
13. http://en.wikipedia.org/wiki/Learning_communities (Accessed March 23, 2009); Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist, 52,* 399–413.
14. http://en.wikipedia.org/wiki/Computer-supported_collaborative_learning (Accessed March 23, 2009); Koschmann, T. D. (1996). CSCL: Theory and practice of an emerging paradigm. In Koschmann, T., Hall, R., & Miyake, N. (Eds.) (2001). *CSCL 2 carrying forward the conversation*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
15. Shell, D. F., Husman, J., Cliffel, D., Nath, I., Sweany, N., & Turner, J. (1997). *Project CIRCLE: Second year evaluation report* (Grant #R215D30195). Washington, DC: U.S. Department of Education: Secretary's Fund for Innovation in Education and Austin, TX: The University of Texas, College of Education, Learning Technology Center.
16. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.
17. Elliot, A., & Harackiewicz, J. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70*(3), 461–475; Lehman, S., Kauffman, D., White, M., Horn, C., & Bruning, R. (2001). Teacher interaction: Motivating at-risk students in web-based high school courses. *Journal of Research on Technology in Education, 33*(5), 45–52.
18. Roeser, R., Midgley, C., & Urdan, T. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology, 88*(3), 408–422.

19. Dweck, C. S., & Leggett, E. S. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.
20. Dupeyrat, C., & Mariné, C. (2005). Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults. *Contemporary Educational Psychology, 30*(1), 43–59.
21. Weiner, B. (2004) Cultural plurality into theoretical unity. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited: Research on sociocultural influences on motivation and learning* (Vol. 4, pp. 13–30). Greenwich, CT: Information Age Publishing; Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology, 71,* 549–570.
22. Weiner, B. (2004). Cultural plurality into theoretical unity. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited: Research on sociocultural influences on motivation and learning* (Vol. 4, pp. 13–30). Greenwich, CT: Information Age Publishing.
23. Dweck, C. S., & Master, A. (2008). Self-theories motivate self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 31–52). New York: Erlbaum/Taylor & Francis Group.
24. Chen, C., & Stevenson, H. W. (2005). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American, and East Asian high school students. *Child Development*, *66*, 1215–1234; Stevenson, H. W., Lee, S., & Chen, C. (1994). Education of gifted and talented students in Mainland China, Taiwan, and Japan. *Journal for the Education of the Gifted*, *17*(*2*), 104–130.
25. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*: New York: Ballantine Books.
26. Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown, and Company.
27. Stevenson, H. W., Lee, S., & Chen, C. (1994). Education of gifted and talented students in Mainland China, Taiwan, and Japan. *Journal for the Education of the Gifted*, *17*(*2*), 104–130.
28. Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263. Dweck, C. S., & Master, A. (2008). Self-theories motivate self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 31–52). New York: Erlbaum/Taylor & Francis Group.
29. Schunk, D. H. (2008). Attributions as motivators of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 245–266). New York: Erlbaum/Taylor & Francis Group.
30. Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370–396.
31. Husman, J., & Hilpert, J. (2007). The intersection of students' perceptions of instrumentality, self-efficacy, and goal orientations in an online mathematics course. *Zeitschrift für Pädagogische Psychologie/ German Journal of Educational Psychology*, *21*, 229–239.
32. Husman, J., & Lens, W. (1999). The role of the future in student motivation. *Educational Psychologist, 34*(2), 113–125; Zimbardo, P. G., & Boyd, J. N. (1999). Putting time in perspective: A valid, reliable individual-differences metric. *Journal of Personality and Social Psychology, 77*, 1271–1288; Husman, J., & Shell, D. F. (2008). Beliefs and perceptions about the future: A measurement of future time perspective. *Learning and Individual Differences, 18*, 166–175.
33. Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*, 954–969.
34. Oyserman, D., Terry, K., & Bybee, D. (2002). A possible selves intervention to enhance school involvement. *Journal of Adolescence, 25*(3), 313–326.
35. Oyserman et al., p. 313.
36. Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive Processes in Self-Affirmation: Intervening to Close the Minority Achievement Gap. *Science, 324*(5925), 400–403.

37. For a recent review on the sources of self-efficacy see Usher, E. L. & Pajares, F. (2008). Sources of self-efficacy in school: A critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796; Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman; Parjares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 111–140). New York: Erlbaum/Taylor & Francis Group.
38. Schunk, D. H. (2008). Attributions as motivators of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 245–266). New York: Erlbaum/Taylor & Francis Group.
39. Zeldin, A., Britner, S., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 48*(9), 1036–1058.
40. Zeldin, A., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal, 37* (1), 215.
41. Wilson, K. M., Trainin, G., & Laughridge, G., Brooks, D., & Wickless, M. (accepted). Our zoo to you: The impact of zoo animals in the classroom on science and literacy concepts in first grade journal writing. *Journal of Early Childhood Literacy*.
42. Tobias, S. (2006). The importance of motivation, metacognition, and help seeking in web-based learning. In H. F. O'Neil & R. S. Perez (Eds.), *Web-based learning: Theory, research, and practice* (pp. 203–220). Mahwah, NJ: Lawrence Erlbaum Associates; Paris, S. G., Yambor, K. M., & Packard, B. W. (1998). Hands-on biology: A museum-school-university partnership for enhancing students' interest and learning in science. *The Elementary School Journal*, 98, 267–288.
43. Cisneros, S. (1992). Eleven. In *Woman hollering creek and other stories.* New York: Knopf; Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology, 92*(2), 331–341.
44. The Oyez Project, FCC v. Pacifica Foundation, 438 U.S. 726 (1978); http://oyez.org/cases/1970-1979/1977/1977_77_528 (Accessed March 23, 2009).

Chapter 9
Efficient Instruction

Cognitive Load

We noted in Chapter 6 that instructional practices developed by cognitive load theorists are fully compatible with the principles of the ULM. Rather than review the details of the extensive instructional literature developed within cognitive load theory (or its equivalent), we refer you to two excellent references. The first is a recent book by Clark, Nguyen, and Sweller, which explicates essentially all of the work published by Sweller and his colleagues, and Mayer and his colleagues.[1] Earlier also, a book was written by Clark and Mayer on the same set of topics.[2]

The CORE Lesson Model

An instructional approach consistent with the ULM principles is the CORE lesson model developed by Calfee and colleagues, which guides teachers as they plan to integrate these fundamental principles into their teaching.[3] CORE is an acronym that stands for CONNECT – ORGANIZE – REFLECT – EXTEND. At the introduction of the lesson and throughout it as necessary, the teacher focuses students' thinking by connecting them to their prior knowledge or experiences either from previous lessons or from commonly held experiences. The teacher guides students in organizing their thinking to build connections with the lesson's information and/or procedures. This is frequently accomplished by teaching with graphic organizers that can be used as discussion or reading comprehension tools. Throughout the lesson, the teacher provides for repetition and connection by asking questions that encourage the students to stop and reflect upon what they are learning in the lesson ("What examples of amphibians should go in this section of the matrix?") and how they are learning the new information or procedures ("Why should we write the common characteristics of the two pre-war periods in the middle portion of the Venn diagram?") that span Bloom's taxonomy. The teachers plan ways to extend students' use of the new learning to other contexts. This portion of the lesson gives students the opportunity to practice applying their new knowledge in varied ways. The teacher asks questions at the close of the lesson that recap the learning

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_9,

and explicitly require the students to think about when they could use their new declarative and procedural knowledge in the future.

Explicit Knowledge Is Teachable; Implicit Knowledge Isn't

Learning is about connections. Explicit mental searching is rather straightforward. You set your mind to a target and then go looking for that target. You almost certainly know whether or not you have ever been in Nome, Alaska. You don't really know how this works. That is, you are not aware of any kind of spreading activation as it takes place. One of us lived in Alaska – but Nome is not connected to the rest of Alaska by roads. So, even when you've lived in Alaska, your mental search for yes/no on having been to Nome usually is a quick search. Since most of us know whether or not we've been in Alaska, the search usually is a quick and accurate one.

Implicit mental searching is quite different; you're not too sure what it is that you are looking for. There is no doubt that much of the real-world problem solving involves implicit rather than explicit searching. An argument in favor of implicit searching as an instructional tool is that it gives us real world practice. Many examples in science and business involve finding information that has similar characteristics to the problem at hand, but is not the same and sometimes very different in appearance.

The instructional question is whether implicit searching is preferable to direct explicit instruction. Learning is about building searchable connections. A teacher or curriculum designer certainly knows more about the relevant knowledge and how it should be organized and connected than a student. Direct instruction based on a good design will certainly build a students' knowledge base faster and build more explicit retrieval cues. As a result, students have more knowledge to search and can explicitly search faster and more accurately. But, even the best planned direct instruction cannot cover all of the possible ways the material could be interconnected or anticipate all of the possible situations in which knowledge might be usefully retrieved through implicit searching for ill-defined problems. Therefore, providing opportunities for students to engage in activities requiring implicit searching can help them build a richer knowledge base with more potential pattern matches.

One can learn without deliberately having a learning goal or consciously paying attention. Episodic memory is effortless memory; it happens. Episodic memory captures space in working memory without conscious effort. Learning by multiple exposures does happen; after a while, we get the sense of things. We also construct novel knowledge in working memory as we are problem solving or thinking. As we discussed in Chapter 4, learning from experience is implicit or sometimes called tacit learning.[4] Knowledge is more difficult to acquire tacitly. As we discussed in Chapter 4, implicit learning from problem solving or critical thinking is uncertain. From an instructional standpoint, tacit learning is more difficult to direct. The teacher or instructional designer can structure the environment in ways that hopefully promote the desired experience, but he/she cannot control whether students are directing their working memory allocation to the knowledge they are supposed to be learning. Mentoring or lengthy apprenticeships are instructional approaches often used. However often we may choose to use the term medical science, much medical practice is learned through residency and fellowship programs in which, what is learned has never successfully been written down. In fact, one of the most often studied areas of learning involves reading x-ray images on the way to becoming a radiologist.[5] Surprisingly, much of a radiologist's learning is tacit learning. Those in training, see a film, and then have an expert interpret that film.

You cannot turn off your working memory. Let's say you are in an apprenticeship, a learning situation. You hear what your master says and see the choices she makes, but she never explains why that choice is made. She might even say something like "I sure hope that works." What you have to process includes the circumstances, what is said, and the seeming decision points that appeared to emerge. There is no apparent rule or decision tree to apply to the situation. It's a bit like learning to bluff when playing poker or making an occasional "psych" bid in playing bridge.[6] Think about it; there can't be a rule for bluffing. If there were, bluffing would never work. If you ever discovered a rule for bluffing at poker, you'd be a fool to tell anyone that rule.

That said, nearly anything that is tacit ultimately could be made explicit. Once it is explicit, it can be taught using more direct instructional approaches, which will allow students to learn more readily. The most efficient way to engage in explicit learning is to use our working memories to think about what we are learning. That is, we attend to what we are learning with the intent of learning it – how it works, what the variables are, and so forth. We repeat; repetition is needed for learning. When we have a chance to think about what we are learning, we can better integrate it with what we already know. That is, we make connections that make sense to us. Good teaching and instruction can explicitly direct students to attend and guide repetition and connection.

Teachers can make it more likely for their students to make these connections through the questions that they ask during and especially after a lesson. Think about this scenario. You are teaching a lesson on the three branches in the American government. After asking the students to brainstorm what they know already about how the government is organized and then giving them the choice of books or websites to read for additional details, you help your students sort out all of this information by developing a chart or grid with them. The chart has headings across the top with the names of the three branches. Down the left-hand column you've written some of the characteristics like: function, duties, people, building where the branch is housed, and one interesting fact. After modeling how to think about the chart and interactively completing a few of the spaces during whole class instruction, you have your students complete the chart in small groups. The next step has the class come together as a whole group to compare/contrast the three branches of government using the information they have recorded on their charts. You extend and apply their learning by asking them to analyze a short scenario about how each branch of government might contribute to the passage or blocking of a proposed law. To close the lesson, you ask two kinds of questions – the product and process questions. Your product question may be as simple as, "Okay, what were we learning about in this lesson? Now, tell your neighbor an interesting fact from the readings or our discussion that was new for you. Sarah, what new information did David learn today? Besides for the quiz tomorrow, why is it important for you to know about these branches of government?" This is followed by your process question, "What did we use to organize our thinking about the three branches of government? Yes! We used a grid. When do you think you could use a grid again to help you think about what you've read? Why would it be helpful then?" When you employ a graphic or visual organizer as a learning/discussion/writing tool, you are making it easier for students to see and make explicit connections between and among ideas. Asking them to make a value judgment pushes them to think about the information more deeply than just the basic knowledge level. The process question pushes your students to think explicitly about transferring a way of thinking about information to different contexts. As soon as we start to mention thinking, we are in the realm of working memory – its capacity and utilization. We've already discussed the notion of cognitive load, and indicated that efficient learning depends upon managing cognitive load. Remember, teachers are working memory managers/trainers/coaches.

This is very much in line with what we learned in the past two decades about teaching how to read. Some students will pick up decoding just by interacting with texts through the scaffolding of more experienced readers (sometimes called "the whole language approach"). This leaves a significant number of students with less than optimal opportunities to learn to decode words, especially if they have limited experiences. At the point of school entry, it is much more efficient and effective to teach the alphabetic system explicitly than to try and supply enough implicit experiences to foster such development. In many ways the implicit approach might actually increase the gap between the more able readers and their struggling peers. Stanovich called it the Matthew Effect in which the rich are getting richer by having more practice and the gap between students increases with time.[7]

Optimal "Difficulty" for New Content

Learning requires connections; we need appropriate prior knowledge to connect to new knowledge. Efficient learning requires a match between the learner's prior knowledge and whatever new content they are dealing with. We may say that the new material is too difficult for the learner based upon her or his current skills. It means that the prior knowledge is either not there, or not chunked in such a way that the new knowledge can be brought into working memory in a way that allows for suitable new connections. Working memory capacity can be used up on small but necessary chunks without leaving enough room for other chunks that are needed to make sense of the new content. This is especially clear when learning a new language.

The other side of this is that, if the new material is too easy, the learner may not pay enough attention – thinking that they already have command of that content.[8] New learning requires working memory allocation, and learners sometimes think they already know something, so they really don't attend to it.

The ULM suggests that for efficient learning to occur, there is an optimal content difficulty for new material that is to be learned. Vygotsky called this the zone of proximal development.[9] What the ULM does is make the requirements more explicit for being in the zone of proximal development. That is, *the material to be learned must fill or nearly fill the learner's working memory without exceeding the learner's working memory capacity*.

In a classroom of twenty to thirty students, it is extremely unlikely that all students will be well described by one and the same zone. This may be true when the content is completely new (no one knows anything) or old (everyone knows the material) but, for most classrooms, learners have different zones – implying the value of small group instructional configurations.

Storage and Retrieval

Learning is a product of working memory allocation. Learning is about storage of new information. We are almost never satisfied with simply asking whether the new knowledge has been stored. It's like asking someone who is completely lost after having a vast rollout of information, "Did you follow that?" and having them respond with a smile and a nod. Teachers have no direct control of their own working memories, let alone those of their students. If a teacher said, "Now store this in your semantic memory" and expected her students to store it, there would be no effect. Teachers often do say things like, "This is important." That's a cue to pay special attention because you may need to use this or, more often, this will be tested.

We satisfy ourselves that students know something through assessment, and we discuss that separately. The ULM suggests that the best results will occur when we practice retrieving in those contexts that matter. Put another way, this suggests that we "teach to the test." Further, it implies that we "test often." If you don't like the context, change the test; that is, change the test that you are teaching to. Make your tests more like the contexts in which retrieval is likely to take place. The Class-Plus Writing Assessment is one such assessment that is designed to determine how well your students are able to compose an informational text under the most supportive conditions.[10] It offers teachers a template for creating a writing assessment that is developmentally appropriate for their class. That's why pilots and physicians train with simulators. Teachers training pilots and physicians strive to make the learning environment as realistic as possible. There are tradeoffs in costs of learning. More resources are allocated to their training of pilots and physicians because "lives are at stake" as the result of the work.

Formative assessment is a good idea. If the assessments reflect the contexts in which the knowledge is to be used, this is nothing more than practice. Frequent formative assessment really should be just practice. High stakes assessment may actually limit student performance because the emotions (anxiety for example) and possible consequences will occupy some of the working memory slots thus inhibiting effective processing.[11]

Notes

1. Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.
2. Clark, R. C., & Mayer, R. E. (2003). *E-learning and the science of instruction*. San Francisco, CA: Jossey-Bass/Pfeiffer.
3. Chambliss, M. J., & Calfee, R. C. (1998). *Textbooks for learning: Nurturing children′s minds.* Malden, MA: Blackwell.
4. Sternberg, R. J., & Horvath, J. A. (Eds.). (1999). *Tacit knowledge in professional practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
5. Lesgold, A., Rubinson, H., Glaser, R., Klopfer, D., Feltovich, P., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. H. Chi, M. J. Farr, & R. Glaser (Eds.), *The nature of expertise* (p. 436). Mahwah, NJ: Lawrence Erlbaum Associates.
6. http://en.wikipedia.org/wiki/Bluffing (Accessed March 23, 2009); http://en.wikipedia.org/wiki/Psychic_bid (Accessed March 23, 2009).
7. Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 340–406.
8. Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23–31.
9. Vygotsky's terminology includes the "zone of proximal development," or ZPD. (Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.)
10. Calfee, R. C., & Wilson, K. M. (2004). A classroom-based writing assessment framework. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders*. New York: Guilford Press.
11. Gimmig, D., Huguet, P., Caverni, J., & Cury, F. (2006). Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence. *Psychonomic Bulletin & Review, 13*(6), 1005–1010.

Chapter 10
Feedback and Assessment

In Chapter 4 entitled "Knowledge", we spent some time discussing what is meant by knowledge. Here are a few things we might ask students to do:

- Recite from memory, the *Pledge of Allegiance*
- Use your own words to explain the statement, "The state at the center of the earth is igneous fusion."
- Find the product of 77 times 88.
- Find the pressure of water vapor inside a 5.0-L vessel that initially is evacuated and maintained at 25°C into which a 2.0-mL sample of liquid water is injected.
- Write a paper describing the emergence and use of wind turbine power generation in the EU.
- Compare and contrast the Japanese and Swiss health care systems with the intent of recommending the adoption of one or the other for the State of California.

These happen to follow a hierarchy according to the *Bloom Taxonomy of Education Objectives*: knowledge, comprehension, application, analysis, synthesis, and evaluation.[1] Although an increase in difficulty is implied, this is not really a part of the taxonomy. Reciting the *Koran* from memory is difficult even though it is a straight knowledge activity; stating how you liked your dinner at the local diner last night is easy even though it is an evaluation.

In classrooms, we ask students to do things. Sometimes we ask them to do things that are tricky. A trick question one of us often uses is to ask about the water vapor pressure that exists after a small amount of water is injected into a large container. The trick is that not all of the water evaporates. Students usually just pull out an equation and proceed to plug-and-chug away without realizing that the physical reality of the situation is quite different from what the equation is meant for. What one needs to do is to look up the pressure in a table. Sometimes we ask them to do things where there is no right answer – like writing a paper or recommending a policy. Regardless of what we ask students to do to demonstrate their knowledge, assessment and specific feedback are critical. From the ULM perspective, any discussion of feedback and assessment should be thought of in terms of the five rules of the ULM outlined earlier. Before discussing these issues, however, we begin with a brief overview of how we define and think about assessment and feedback.

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_10,

Assessment

Classroom teachers assess students all the time. We give tests, grade homework, and observe students working on group projects. From this perspective, assessment is simply a tool we use to determine how much learning has taken place. From the ULM perspective, *assessment is a tool that we should use to determine how close students are to reaching their goal*. We can do this in a number of ways. What may be the most important issue, however, is not what kinds of assessments we use, but rather what we are doing with them.

The ULM argues that the goal of most learning situations should be mastery, long term retention, and the ability to transfer information to other contexts or situations. As we saw in Chapter 4, if we want to promote mastery, then we need to provide opportunities for practice in varied contexts. Another way to say this is that if we want students to pursue mastery, then they need (1) a goal, (2) opportunities for practice with assessment, and (3) significant feedback on their progress toward that goal. If we also want to promote long-term retention, then students should do things that require long-term retention. From the ULM perspective, the most appropriate way to accomplish both mastery and promote long-term retention is to provide opportunities for regular "formative" assessment and a more formal cumulative or "summative" assessment at the end. Formative assessment is used during learning. It includes any number of activities in which students can receive feedback on their progress. This may include (but is not limited to) questions posed by the teacher or classmates during instruction, in-class activities, homework, informal (and formal) teacher observations of student progress, student reflections, quizzes and exams during the semester, and class projects. The key feature of formative assessment is that it provides teachers with an opportunity to assess student progress toward the goal together with an opportunity to provide students with feedback about their progress including recommendations for improvement.

In contrast to formative assessment, a cumulative or "summative" assessment should be designed to allow teachers to assess – at the end of instruction – the extent to which a student achieved mastery and long-term retention. Additionally, summative assessment should provide students with an analysis of whether or not they have reached their instructional goals. Summative assessments are usually more "formal" than formative assessments. In many cases, a cumulative assessment regarding long-term retention is built directly into the content students are learning. In a typical math class, for example, what you learn in Chapter 4 will likely be used to complete more complex tasks in Chapter 6; so long-term retention is a necessary prerequisite for future success. This may not be as obvious in other areas. For example, in a psychology course, what you learn about Freud may, or may not be, a prerequisite to understating Piaget's theory. Whether or not the subject has some cumulative long-term retention requirements, the question is whether or not students learn in that way. From the ULM perspective, a well-designed summative assessment encourages long-term retention.

Feedback

No matter how we assess students, the information we provide them about their performance on that assessment – the feedback – may be what matters most. Feedback is information provided to a student by a teacher about that student's performance, progress, or understanding. We know that some types of feedback impact learners more or less than other types of feedback. In general, feedback about the process students are using tends to be more effective than feedback about the product. Therefore, simply grading homework with a percentage score (e.g., 80%) or a summative comment (e.g., great work) may not be sufficient feedback for the most effective learning. Rather, substantive feedback about the process students are engaging in seems to be the most effective form of feedback. We can say three things about feedback. First, effective feedback is a required part of the learning process. Second, effective feedback occurs at multiple points during learning. Third, but perhaps most important, effective feedback prompts students to generate their own feedback. When a teacher provides feedback to a student by asking him to think about how he reached a particular answer or response, he must think about what he did and thus generate a personal assessment of his progress. In short, from the ULM perspective, effective feedback prompts students to allocate working memory space to self-monitoring or self-regulatory activities. This is not easy. Developing effective feedback can be very time consuming and requires effort on the part of the teacher. Our intent here is to use the ULM to guide the way we provide feedback. We begin by looking at how assessment and feedback fit within the ULM rules.

New Learning Requires Attention

The first ULM rule is that *teaching is about getting students to "attend" to things*. Perhaps the most basic function that assessment performs is telling students what to pay attention to. How often have we heard students ask "Is it going to be on the test?" or say "If it isn't on the test, I am going to skip it." For better or worse, students base what they are going to attend to and going to put forth effort to learn on what they think will be tested. While this is often viewed as bad and "teaching to the test" is seen as lowering the quality of learning, assessment and testing are not inherently good or bad. Good assessment takes advantage of the privileged role that testing has in directing attention by assessing what is important. If assessment is in line with the goals of instruction, then the assessments will reinforce the learning goals of the class. It is only when assessments are not aligned with learning goals and test peripheral or non-important information that they become problematic and dysfunctional for learning. Similarly, feedback performs a necessary function by helping students direct their attention to the proper instructional materials or learning activities and reinforcing what is and isn't important.

The attention directing roles of assessment and feedback mean that careful attention must be paid to creating high-quality tests and other assessments that are clearly and tightly aligned with the learning goals of the class. A classic way of doing this is through a *table of specifications*, which is a grid that crosses learning objectives/goals with assessment items and activities. Excellent guidance and examples of these can be found elsewhere.[2] Also, feedback must be well thought out and focused on specific information that needs to be attended.

Learning Requires Repetition

The second ULM rule is learning requires repetition. Repetition is accomplished by retrieval of knowledge from long-term memory into working memory. Retrieval means something going out. Assessments require demonstration of knowledge and hence require that knowledge to be retrieved. The act of retrieving knowledge for an assessment not only demonstrates that it was learned, but increases the strength of that knowledge in memory. A recent and very interesting study under controlled conditions provided a rather remarkable outcome that illustrates this point.[3] A large group of students was divided into four groups studying Swahili vocabulary. One group studied and tested only those words they had previously not learned. A second group studied and tested all words. A third group studied only words they felt unfamiliar with, but tested all. A fourth group studied all but tested only those found to be unfamiliar. "Repeated studying after learning had no effect on delayed recall, but repeated testing produced a large positive effect. In addition, students' predictions of their performance were uncorrelated with actual performance. The results demonstrate the critical role of retrieval practice in consolidating learning and show that even university students seem unaware of this fact."

Assessment performs an important role in strengthening what has been learned. Assessment provides opportunity for practice and success comes to those who practice, practice, practice. We may sometimes hear a student say "I studied and knew the material, but I just couldn't bring it out on the test." What this likely means in ULM terms is that this student hasn't had enough repetitions of the knowledge to make the chunk strong enough to retrieve. Assessment can provide that repetition. Assessments also provide opportunity for corrective feedback to help refine knowledge in memory chunks. So, more specifically, we believe that success comes to those who practice, assess, get feedback, practice, assess, get feedback, and so on. Without assessment, it is impossible to provide guidance to the repetition necessary for accurate chunk formation.

Learning Is About Connections

Knowledge is about creating a connected pattern of neurons. Higher, more complicated forms of learning, like concept formation or skill development, are essentially about forming a series of complex connections among discrete ideas. Students make connections between the concepts of shapes and sides to understand the

Pythagorean Theorem or among historical issues, a belief in states' rights and the Emancipation Proclamation as they generate an understanding of issues that led to the Civil War.

Assessment fosters creating these connected patterns by requiring students to connect and combine information and skills they are learning. Asking students to compare and contrast discrete concepts, solve problems, transfer knowledge to new contexts or situations, write essays, demonstrate skills, and other forms of "open-ended" assessments allow students to engage in the problem solving and critical thinking that underlies the construction of larger knowledge structures in memory and the creation of new connections. Assessments that require making connections will foster knowledge connections. These assessments also provide opportunities to observe the constructions and connections students are making so that effective feedback can be provided to foster productive constructions and avoid misconceptions. Open-ended assessments provide rich opportunities for mentoring and guidance and more indirect instructional feedback methods like Socratic dialog.

Teaching to the Test

Much talk centers on the phrase, "Teaching to the test." This has been especially true since enactment of the No Child Left Behind legislation. Many, if not most people, seem to react negatively to this sound bite. It connotes very restricted learning. There was an advertisement on local television in 2008 for learning centers intended to help struggling students. The ad showed a student who one assumes had learning issues and was helped by the center. She recited the last names of the presidents: "Washington, Adams, . . ., Bush, Clinton, Bush." Obviously this was impressive. To accomplish this required effort and, because of that, required motivation. It's a task that was doable. Success with a task like this might increase self-efficacy for learning, especially if parents were to ask the child to demonstrate this skill to friends. (OK, did you remember that Fillmore was a president?) On the other hand, how useful was it? It's not that knowing this is either bad or wasted; it's that this knowledge alone is not particularly useful. Moreover, given just a list, the learning strategy involved is straightforward: rehearse, rehearse, rehearse. If the test consists of recalling lists of presidents' names, then perhaps what is being tested is not in the overall best interests of national security.

Nevertheless, we recommend that teachers re-contextualize their use of the phrase, "teaching to the test." What we would like to do, under the best of teaching circumstances, is to provide opportunity for transfer to the contexts most like those in which we believe that the knowledge being learned is going to be used. If this is the case, then the "tests" probably need to be changed to better reflect the ultimate intended contexts. While this is quite difficult to achieve, it should remain a goal.

Many have embraced Wiggins & McTighe's *Understanding by Design*.[4] This approach logically advocates deciding what you want known, then deciding how you will know it is known, and then creating instructional materials that will lead to the appropriate learning. At the end of the day, this is really just a repackaging of teaching to the test. Think of it this way: If it isn't something you wanted learned, why was it on the test? If it *is* something you want learned, why *isn't* it on the test? We emphasized the use of tests to reinforce classroom learning goals in our discussion of how assessments focus attention.

In the United States as well as most other countries, the purpose of compulsory education is to "centralize" those aspects of knowledge and intellectual skill thought to apply to all or most vocations. As a result, reading and arithmetic are taught in schools. In contrast, surgery is taught in surgical rotations in medical or veterinary school, heat transfer in advanced engineering, and steamfitting in apprenticeship programs. Both the reading and arithmetic skills used by surgeons, engineers, and steamfitters are different. However difficult and impractical, it would be better and would lead to better learning outcomes if reading and arithmetic could be contextualized at those levels. That is, if we could say: "This is what a steamfitter needs to know about geometry." On the other hand, a great strength of education is that, if you are a steamfitter and want to become a heat-transfer engineer, you don't have to repeat your K-12 education to make the change.

High-Stakes Testing Versus Feedback

One certain impact of the U.S. No Child Left Behind legislation has been an emphasis on testing, especially what one might call high-stakes testing. High-stakes testing is not likely to go away any time soon, nor should it. Stakes in tests like SAT, GMAT, or MCAT are high. These stakes pale in comparison with, say, the boarding exams taken by those seeking certification and/or licensure in medical specialties or those taking "bar" examinations to practice law or those becoming commercial airline pilots. There *are* times when assessment for the purpose of measuring and documenting knowledge is appropriate.

There are several problems with high-stakes testing. These tests are expensive and complex to develop, organize, and administer. For this reason, they often are not available "on demand." Emotions run high for these tests, and adverse effects on performances can result. Students often "choke under pressure."[5] Perhaps most importantly, high-stakes tests, as currently employed, often do not facilitate a focus on either mastery or deep and meaningful learning. Rather, high-stakes tests seem to facilitate a focus on learning "just enough" to perform well on the test. From the perspective of the ULM, high-stakes tests can create goals for performance or task rather than for learning. This should lead us to the question of what is the alternative.

As we have discussed, providing performance-related feedback is one alternative that is an effective way of bringing about learning. As students engage in a learning task, teachers can monitor their progress and provide feedback to students about not only what they are doing well and not so well, but also about how they are going about doing it. This ongoing feedback is helpful because it gives students information about what they are doing before they have automated it. At the same time, it gives teachers very powerful ways to decide about a learner's knowledge.

Modern computer-based homework systems and other electronic practice systems offer excellent means of providing "formative" assessments of learning. In these systems, students have an opportunity to submit their homework electronically and to receive feedback on their work as they are working on it. At a minimum, these systems can provide feedback to students before they have automated a new task. Remember, it is much easier to correct someone before they have automated a task. If a teacher waits until his student has automated a task, it becomes much more difficult to correct errors. Earlier on p. 48, we discussed how hard it was for Tiger Woods to change his swing.[6]

Praise Versus Encouragement

For two decades, some in the literature have suggested distinguishing praise and encouragement in the classroom.[7] Praise involves saying nice, pleasant things without indicating why that compliment is being given. Encouragement spells out rather explicitly why the praise is being offered. From the perspective of the ULM, it is a matter of explicitly connecting the praise to whatever action is being praised. These connections permit a teacher to encourage mastery rather than performance. Consider Table 10.1.

Table 10.1 Examples of feedback

OK	Better	Much better
Nice job, Fred.	You seem to be working very hard on that, Fred.	Nice work, Fred. I see a lot of progress in lab reports. Your hard work is really paying off.
Well done, Mary.	That first draft looks good so far, Mary.	This is a good start Mary. You have done a nice job describing the main problem. Now, let's see what we can do to improve on this even more.
Very good, Sally.	Your data analysis seems quite clever, Sally.	I am impressed with how far you have come in your data analytic strategy Sally. The way you handled the missing data was very creative.
I really like that sound, Jamal.	You must have been practicing quite a bit last week, Jamal.	I like what I am hearing Jamal. You are really benefiting from all the practice you have been putting in. Your rhythm is really strong.
I'm so glad you are in my class, George.	Thank you for helping others, George.	George, You have really made some important contributions to the class. I know your classmates appreciate how much you have helped on the story problems.

Scaffold Learning by Responding to Outputs

Essentially all teachers advocate some form of active learning – in which the learner is engaged in his or her own learning processes. Active learning is difficult to define or describe. In the ULM, active learning is motivated working memory allocation. Of course this isn't really observable, so it may be difficult for someone like a teacher to determine if learners are actively engaged with new content. One fairly certain way to detect active learning, however, is to interact with learners by providing performance-related feedback. During instruction, an emphasis on eliciting outputs from students and responding to those outputs with suggestions for improvements is likely to be the most efficient way to keep students learning "actively" and ultimately bring about new learning. Perhaps you can see why feedback must be focused on the processes students used, rather than simply the outcome.

Learning is not just a "garbage in, garbage out" process. While much of what we teach is explicit and can be learned as declarative knowledge, success in life nearly always depends to some degree on developing skills for using knowledge – creating procedures. If teachers want that to happen, then ways of using knowledge and creating skilled behaviors are what should be practiced and tested.

Responding with feedback to student outputs is especially critical for proceduralization. We discussed how procedural knowledge was created in Chapter 4. Recall that procedural knowledge is based in the results of action; success in achieving a goal. How do students know if their results are correct? How do they know whether they are building optimal procedures? The only way they can know is through feedback.

Declarative knowledge is typically right or wrong; you either know it and have strengthened it enough through repetition to recall it on a test or in another setting, or you don't know it. Procedural knowledge, on the other hand, is often not clearly right or wrong. There can be many ways to do something, many approaches to solving the math problem, or many options for conducting a scientific inquiry. Helping students sort out these possibilities and strengthen the best approaches is what teachers can do with feedback. This is what scaffolding student learning is about. This is also the heart of what Ericsson called deliberate practice with a mentor.[8]

Assessments should allow students to use knowledge in ways that provide observable outputs. From those outputs, teachers can give feedback that effectively scaffolds for the student. Use of knowledge requires assessments that go beyond just statement of the knowledge. We might want to teach a rule such as: In the periodic table of the chemical elements, atom size increases going down a group. This could be tested with something like: state the trend in size for groups in the chemical periodic table. This would be a direct test of declarative knowledge. The same rule might be tested as follows – select the element with the largest atoms from among: lithium, sodium, and potassium. In this case, the learner must decide that all of these elements are in the same chemical group, that potassium is the lowest in that group, and that potassium would have the largest atoms. This learning involves acquiring a procedure that is a decision rule for determining the largest atom in a group, given the knowledge that atom size increases going down a group. This procedure would apply to any combination of atoms in the periodic table. Besides allowing a teacher to know whether the student knows the rule, this kind of assessment allows feedback for developing a procedural application of the rule.

Notes

1. Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals* . New York: D. McKay. This has been revised slightly to remember, understand, apply, analyze, evaluate, and create as found in: Anderson, L., & Krathwohl, D. (Eds.). (2001). *A Taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives, complete edition*. Columbus: Merrill. The net effect of this would be to interchange the fifth and sixth items of our example. In this example, we believe that the last item is the most challenging of those listed. A common confusion about the taxonomy is that it corresponds to difficulty. Naming the presidents or reciting the Koran from memory both are the "lowest" taxonomic levels. Responding to, "How was dinner last night?" is an evaluation.
2. Kubiszyn, T., & Borich. G. (2007) *Educational testing and measurement: Classroom application and practice* (8th edn.). Hoboken, NJ: Wiley/Josey Bass Education.
3. Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968.
4. http://en.wikipedia.org/wiki/Understanding_by_Design (Accessed March 23, 2009); Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (2nd edn.). Alexandria, VA: Association of Supervision and Curriculum Development.
5. Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail. Working memory and "choking under pressure" in math. *Psychological Science, 16*(2), 101–105.
6. http://www.oneplanegolfswing.com/oneplanemembers/Tour_Pros/Tiger-Woods/index.jsp (Acessed March 23, 2009).
7. Hitz, R., & Driscoll, A. (1988). Praise or encouragement? New insights into praise: Implications for early childhood teachers. *Young Children, 43*(5), 6–13.
8. Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Chapter 11
A Focus on Thinking

Check in the index of a book on cognitive psychology or instructional design and you're not likely to find an entry for "think."[1] Ask a college science teacher what their goals for students include and you almost always hear "teaching them how to think." Look up "thinking" in *Wikipedia* and you find "... allows beings to model the world and to deal with it effectively according to their objectives, plans, ends and desires." Your dictionary will likely express several different definitions, all of which you'll have heard and be able to relate to.

What is thinking in the ULM? We have alluded to critical thinking and problem solving in Chapter 4 as continued search in working memory for new sensory input (new information) and restructuring and transforming available sensory input into different configurations for the purpose of creating a pattern match. We want thinking of this sort in our classroom when we want students to use what they have learned to solve problems or answer questions. We would like them to be able to "think" about their answers beyond simple rote regurgitation.

We also want students to think about what they are learning. Thinking in this context means something a little different. It means making sense of what they are learning. Students need to build their knowledge by transforming it and connecting it to existing knowledge. Creating pattern matches that retrieve prior knowledge to help understand what is being learned and place it in context.

Thinking involves working memory. As we learn new things, the intrinsic load of the new material may be such that we can't do too much beyond the simplest processing with what currently is active in working memory. While we can successfully learn how to do what we need to do at the moment, we don't really get the chance to see why we are doing it or how what we are doing fits in with other things that we know.[2]

Keep in mind three things about encouraging thinking: learning requires repetition, learning is about connections, and learning can either be effortless or require effort. At the end of a lesson, it is very reasonable to review what has been covered and to connect that new learning to other things as described in several examples we have already mentioned. This fits well the rules of repetition and connection. The context often is a casual one, and thus this learning is stored in episodic memory;

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_11,

it is effortless learning. If the learning is not emphasized in some way, it is subject to the vagaries of episodic memory. Effort is almost certain to be required to help move this from episodic to semantic memory. We'd like to suggest some more explicit things worthy of your consideration that are intended to help the conversion from episodic to semantic memory.

Surprisingly, in the development of cognitive load theory that has been a principal area for study in instructional design, the consideration of germane load – which we could describe unceremoniously as thinking – did not emerge until the last decade. Therefore, the literature in this area is thin. We'll remind you of some important aspects of content-specific thinking, and then move to generalities.

Critical thinking and problem solving in the ULM involve both continued search for new sensory input (new information) and restructuring and transforming available sensory input into different configurations for the purpose of achieving a pattern match for LTM retrieval.

Content-Specific Thinking

Every discipline has its own way of thinking. Doctors think like doctors, lawyers think like lawyers, police officers think like police officers, and so on. If you are teaching a discipline, as in most high school and college teaching, you need to become familiar with the ways of thinking in that discipline and try to convey that to your students.

In the chemistry of carbon compounds, for example, we expect each carbon atom in a molecule to have four chemical bonds or connections to other atoms. Organic chemists automatically check for this. Whenever a molecule is represented with a carbon atom having either more or fewer than four bonds, we either predict special reactivity for that atom or suspect that an error has been made in the representation. In mathematics, we learn that multiplying an entity in an equation by one (unity) leaves that entity unchanged. Sometimes we get to be very clever in the way we write one (unity). For example, we might write one as $((\sin x)/(\sin x))$. When studying herd animals, we expect there to be an alpha animal – a leader from whom most or all other animals take their lead. When studying economics, we expect savings to equal spending. When studying psychology, we expect behavior that is rewarded to occur more frequently. When studying spread of viral disease, we need to consider airborne transmission as a mechanism. We expect the eating habits of children to be more like those of their parents than anyone else.

Each discipline has its rules, and sometimes these can be very complex. Sometimes the rules can be simple but not easily verified. Given only four colors, one can uniquely color a map so that no two adjacent entities have the same color – the so-called four-color theorem. Though first proposed in 1852, this theorem was not "proven" until 1976, a proof that has led to other improvements and proof strategies.[3]

Have Students Anticipate (Expectancy-Driven Methods)

In an expectancy-driven method, the learner anticipates an outcome and then matches that expectation against feedback about the outcome. Renkl proposed this approach in studying probability.[4] A similar result was found when asking learners to anticipate the functioning of a machine before seeing an animation of that device.[5] We think that the *discrepant event* used in science education generates situational interest. In this sequence the teacher indicates that she will do xyz and asks the students to predict the outcome. The teacher then does xyz, and the students get to compare their predictions with the actual outcome. As the label suggests, the events are selected such that the predications most often prove wrong. That is, events are selected such that a discrepancy between the likely prediction and the actual outcome is anticipated. For example, a weight is hung from a rubber band, and the teacher says: "I'm going to heat this rubber band with hot air from this hair dryer. What will happen?" The students most often predict that the rubber band will soften and the weight will go lower (i.e., the band will stretch). In fact, the rubber band contracts and the weight rises – unless you heat it enough to destroy it. This effect can be built into more complex if classic demonstrations.[6]

From the perspective of the ULM, the expectancy-driven method provides opportunity for new chunk building and new connections. Discrepant events tend to attract attention, so they can focus students' attention on the material. As students encounter the discrepancy between their prediction and the outcome, they have opportunity to examine and transform the connections in their existing knowledge that were incorrect in light of the new information provided by the discrepant outcome. The discrepancy serves to weaken existing elements of the chunks inconsistent with the new information and strengthen new connections that incorporate the experienced outcomes. Follow-up discussion can help further weaken incorrect knowledge and strengthen the new connections.

Everyone reading this book is most likely using an expectancy-driven method for constructing meaning. Proficient readers automatically make predictions about what they will be reading. They do this by looking at the title of the book. They predict what a section will be about from the chapter titles, subheadings, and any illustrations the author has included. As they read, good readers confirm or disconfirm their predictions, sometimes causing them to reread to be sure they read a sentence or passage correctly. This recursive act is not automatic for beginning readers or students of any age who struggle with reading. These children need explicit instruction and practice in how to predict as they read. As novice readers become more facile in using prediction, they begin to focus their attention better on monitoring their comprehension and resolving prediction discrepancies when they arise.

Teachers Create Sub-goals (Parse the Intrinsic Load)

What can the teacher, trainer, or instructional designer do when a learning task is very difficult? The standard approach is to divide the task into subtasks so that the learner's working memory capacity is not overloaded with new input; in cognitive load theory terms, reduce the intrinsic load of each subtask so that it is within the learner's grasp. In ULM terms, insure that the amount of material to be learned doesn't exceed the learners' working memory capacity. As we have discussed, the learner's working memory capacity is increased by prior knowledge, so more knowledgeable students can handle more input and deal with more complex manipulations of information. Vygotsky noted this need to adapt learning materials to the learners' capacity as matching the learner's zone of proximal development. Catrambone was among those to detail this approach to teaching complex problem solving.[7] This approach also accounts for some of the success found when using worked-examples during instruction. When a student views (or memorizes) a solved problem, the steps can be accessed as sub-goals and used to reduce the overall intrinsic load.

Sub-goals are potentially important for a different reason. When a novice learner is working near mental capacity, there is little available working memory capacity to "think" about the problem being worked on, to see how the current step and goal fits with other steps and goals. Winne expressed this in terms of the load placed by self-regulation upon some beginning learners. If learners have to manage too much of the learning environment themselves, they may be unable to devote sufficient capacity to creating the transformations and connections necessary for effective learning.

Remove the Scaffolding

As students become better problem solvers in a given area such as learning to read, their thinking improves when instructional scaffolding is lowered or removed. In the 1980s Pearson and Gallagher called this the Gradual Release of Responsibility Model of Instruction.[8] The term from behavioral approaches to this instructional strategy for enhancing thinking is fading.[9] Fading, or gradual release, is one of the most powerful techniques for helping learners decide when to apply particular strategies. Research has shown a variety of instructional choices that can support and advance learning more quickly than others. Reciprocal teaching, where individuals take on teaching roles within small groups, appears unusually effective.[10]

Have Students Imagine Outcomes

There are two ways to envision some sort of psychomotor process: where you want to be, or how you are going to get there. For example, think about teaching someone to throw darts. In the "internal" approach, you ask them to concentrate on holding and manipulating the dart. In the "external" approach, you ask them to focus on the target and do what it takes to strike the center of the target. A review concludes that numerous studies "... provide converging evidence that an external focus of attention (i.e., focus on the movement effect) is more effective than an internal focus (i.e., focus on the movements themselves)."[11] In a fashion similar to that found for improving psychomotor skills, similar outcomes are found for cognitive skills, but the results are mixed.[12] For students who have the schema in place, imagining works successfully. For those without the schema, imagining does not work so well. This seems to be the same load issue talked about by Winne, namely, when a novice student is near working memory capacity, there's no extra capacity for the kind of thinking that imagination requires. Once past this, however, imagination becomes a useful thinking tool that enhances learning performance.

Accommodate Cognitive Artifacts

American resurgence of interest in science teaching in the 1950s, following closely the launch of Sputnik, led to changes that included emphasis on calculations. Slide rules were devices required of science and engineering students; a slide rule case dangling from one's belt was a give-away about his or her vocation. Electric calculators were available; access to a Monroe calculator was a status symbol. An Wang, founder of Wang Laboratories, developed an electronic calculator. All of this change was dwarfed in 1972 by the emergence of the HP-35 from Hewlett Packard, the world's first "pocket calculator." The science education literature started to fill with papers about the use of calculators – and whether such usage was "fair." Soon calculator ownership became expected of students as prices dropped to the same level as textbooks. Today, high quality graphing calculators are a bit less expensive than most science textbooks!

Something else was happening during the post-Sputnik era: science curricula came to expect very extensive calculations from those who would be labeled "successful" students. In the late twentieth century, it became clear that students were able to compute quantities accurately without having any contextual understanding of what those numbers meant. Quantitative approaches were labeled as algorithmic; algorithmic became a dirty word.

Donald Norman described the tools we use to make our cognitive lives easier *cognitive artifacts*.[13] He makes the important point that these tools change the task. There was a time a few decades ago when a graduating engineer could make a living by solving differential equations. While the task remains important, the procedures have changed so much that having this as a near single skill set is of extremely limited value. Tools such as *Mathematica* enable performing such tasks in a way that has led to significant raising of the intellectual bar.[14] Salomon noted that technology tools produce two effects.

Like Norman's cognitve artifacts, there are effects with technology where the tool allows increased performance. The expertise of the human – tool system is greater than the expertise of either the person or the tool individually. Then there are effects of technology or how the technology impacts the person using it. Something like a computer based writing assistant with an idea generator and word processor can have an effect with a student allowing them to produce a better essay. Also, there is the potential that the interaction with the writing assistance can have an effect of helping the student become a better writer. Salomon thought that for this to happen the tool had to activate higher order thinking or provide a model of higher order skill that the learner could learn.[15]

Experts Practice Deliberately

What do experts do that make them stay experts? Among other things, experts practice. Experts develop routines.[16]

We often hear it said that some people are invested with special talents that are unique. Perhaps. But as we've already noted, talent is probably a word that a teacher wants to expunge. If you embrace this term, then it is better to speak of talent as something one acquires through life, rather than something one has from birth.

When experts are studied, several consistent observations are made. Experts usually have external support – from the likes of family and/or friends.[17] They have access to prior expertise through materials such as those found in a library. They have a mentor or coach. Perhaps most important, they engage in what has been termed *deliberate practice*.[18] Two renowned athletes, Michael Jordan and Tiger Woods, are famous for their dedication to practice. When athletes practice, it's usually obvious – one can watch them! Not so with writers or surgeons or plumbers.

What's the difference between ordinary practice and deliberate practice? In ordinary practice, you are trying to automate some processes; in deliberate practice, you are processing in working memory with the intent either of developing or changing a process. In deliberate practice, you are thinking about some details of what you are doing, possibly with the intent of changing them. We've mentioned more than once the lengths Tiger Woods went to when "changing his swing."[19]

A problem with changing processes that are automated is that they do not come into working memory and take up space. To modify an automated process, we must somehow consciously bring it into working memory where we can work on some aspect of that process. During much routine work, we really don't want to be thinking about what we're doing, especially when we are experts. The expert urologist performs a nerve-sparing prostatectomy and needs to seek and dissect around the neurovascular bundle. She can't be thinking about "how to hold the knife" or "how to control the robot." Before being called a surgeon, long before they are experts, they *do* have to think about how to hold the knife! That's a time when they engage in deliberate practice. If something causes them to think that a skill is slipping – perhaps holding the knife – then they need to return to practice that skill.

Reading is a skill that essentially everyone needs. Even the best readers usually can get better through deliberate practice. If you are reading this book, your decoding almost certainly is automated. Suppose we taught you a new reading comprehension strategy that involved asking and answering questions about the relationships between theory and practice. Certainly, you would have to think about your reading as you learned when and how to use the new strategy. If your reading comprehension stayed automatic and problem free, how would you practice this new strategy and know when to use it? Once you begin to master the new strategy, then you can stop thinking about your automated reading skills. One way to teach struggling readers is to focus on their procedural knowledge of comprehension strategies. More specifically, to discuss what good readers do. Researchers have discovered the practices of good (or expert) readers. The knowledge gained from understanding the automatic procedures used by expert readers is turned into a blueprint for practicing with less able readers and leads to gains in their reading ability.[20]

Conceptual Change

Conceptual change receives unusual attention in education for two reasons. First, most education and especially science education is directed at students' gaining conceptual understanding of the material rather than simple rote factual memorizing. Second, many teachers, especially science teachers, are concerned that students come to their classrooms with misconceptions they have developed from their everyday experience, and teachers are often faced with trying to change these pre-existing concepts. Students go to lengths to hang on to existing notions rather than change.[21] Many times learners confronted with information inconsistent with their current view of the concept ignore the new information or "compartmentalize" the new and old knowledge in separate chunks by keeping mental notes of the sort "but X behaves like Y and not Z in this case." In other words, it's like: "Note to self – in school, say Y instead of Z or you'll be wrong."

Considerable empirical and theoretical work has been directed at attempting to explain conceptual change as evidenced by a number of recent reviews and theory papers.[22] In Chapter 4, we extensively discussed how concepts are learned through chunking. In the ULM, concepts are produced by strengthening of the core characteristics of a knowledge chunk through repetition according to the law of large numbers. So the strength of a concept is a function of how often it has been repeated. Given that chunks/concepts extracted from everyday experience often have had extensive repetition, it is not surprising from the perspective of the ULM that students hang on to existing knowledge even if from a technical standpoint it is a misperception.[23] Like trying to change automatized procedural knowledge or a habitual behavior, confronting and changing highly strengthened pre-existing declarative knowledge chunks is difficult.

As we discussed in Chapter 4, changing automatized procedural knowledge requires "deliberate practice;" bringing the automatized procedure into working memory so that attention can be focused on the procedure. Confronting pre-existing conceptual knowledge is not quite analogous because concepts as declarative knowledge are always retrieved into working memory; there is nothing equivalent to automatized procedural knowledge that bypasses working memory. Getting the misperception into students' "attention" is not the problem.

By the working memory rule of association, when two separate pieces of knowledge are in working memory, they will be connected. So, when a student has retrieved a pre-existing chunk and they are presented with new knowledge while this chunk is in working memory, the new knowledge will be connected to the old.

By the rule of repetition, however, unless this new knowledge is repeated, it will quickly weaken in strength relative to the existing knowledge. If the new knowledge is mutually exclusive (i.e., the old and the new contradict each other), then either the new knowledge will be purged through active inhibition or the new knowledge will be separated into a new chunk. Think of the formation of separate dog and cat chunks out of "animal" as conflicting information about dogs and cats increases. These processes account for both the observations that students simply ignore conflicting conceptual information and that students "compartmentalize" information into broad categories of school and real world.

Like other theories of conceptual change, the ULM unfortunately does not provide an "easy" answer to how to confront students' misperceptions and replace them with more appropriate conceptual understanding. As discussed in Chapter 4, providing best exemplars of concepts and doing instruction that focuses student attention on the salient core knowledge components of the concept will build accurate concept chunks faster and more accurately. In the early grades, maybe the best that a teacher can do is to get students to acquire an accurate "alternative" conceptual chunk, even if it is compartmentalized to school. As students progress through their schooling, there will be more opportunities for the accurate conceptual knowledge to be encountered repeatedly. Hopefully this can strengthen the correct formal knowledge enough for it to ultimately replace the old naïve concept.

As we discussed previously, discrepant events might be used to highlight the differences between the alternative concepts. Certainly the more times the students bring the two contradictory concepts into working memory together, the more rapidly the repetition based concept formation processes will work. Any instructional activities from demonstrations to hands-on experiments or exercises that will make students "think" about the knowledge and any contradictions that need resolution will increase the number of transformations, repetitions, and new connections that happen. But there is no certainty that these methods will lead to replacement of the old chunk rather than compartmentalization.

Notes

1. For example, see Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th edn.). New York: Worth Publishers; or Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.
2. Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist, 30*, 173–188.
3. http://en.wikipedia.org/wiki/Four_color_theorem (Accessed March 23, 2009).
4. Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*(1), 1–29.
5. Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21*(4), 209–249.
6. http://demo.physics.uiuc.edu/lectdemo/scripts/demo_descript.idc?DemoID=578 (Accessed March 23, 2009).
7. Catrambone, R. (1998). The subgoal learning model: creating better examples so that students can solve novel problems. *Journal of Experimental Psychology. General, 127*, 355–376.

8. Pearson, P. D., & Gallagher, M.C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology, 8(3),* 317–344.
9. Reisslein, J., Reisslein, M., & Seeling, P. (2005, October 19–22, 2005). WIP: Effectiveness of worked examples and fading in introductory electrical circuit analysis for learners of different ability levels. Paper presented at the Frontiers in Education, 2005. FIE '05. Proceedings 35th Annual Conference, Indianapolis.
10. Palincsar, A. S., & Brown, A. (1984). Reciprocal teaching of comprehension fostering and monitoring activities. *Cognition and Instruction, 1*, 117–175; Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229–270). Cambridge, MA: The MIT Press.
11. Wulf, G. (2007). Attentional focus and motor learning: A review of 10 years of research. *Gabriele Wulf on attentional focus and motor learning. E-Journal Bewegung und Training, 1*, 4–14.
12. Cooper, G., Tindall-Ford, S., Chandler, P., & Sweller, J. (2001). Learning by imagining. *Journal of Experimental Psychology Applied, 7*(1), 68–82.
13. Norman, D. A. (1991). Cognitive artifacts. In J. M. Carroll (Ed.), *Designing interaction* (pp. 17–38). Cambridge: Cambridge University Press.
14. http://www.wolfram.com/products/mathematica/index.html (Accessed March 23, 2009).
15. Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, *20*(3), 2–9.
16. Kiewra, K., & Creswell, J. (2000). Conversations with three highly productive educational psychologists: Richard Anderson, Richard Mayer, and Michael Pressley. *Educational Psychology Review, 12*(1), 135–161.
17. Gardner, H. (1998). *Extraordinary minds*. New York: BasicBooks; Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown, and Company; Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*. New York: Ballantine Books.
18. Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Rev*iew, *100*(3), 363–406; Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). New York: Cambridge University Press.
19. http://www.oneplanegolfswing.com/oneplanemembers/Tour_Pros/Tiger-Woods/index.jsp (Accessed March 23, 2009).
20. Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*, 1239–1252; Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. 5, pp. 89–129). Greenwich, CT: JAI Press.
21. Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1–50.
22. Vosniadou, S., Baltas, A., & Xenia, V. (Eds.). (2007). *Reframing the conceptual change approach in learning and instruction*. New York: Elsevier Science.; Ohlsson, S. (2009). Resubsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist, 44*(1), 20–40; the entire 1st issue of volume 44 of *Educational Psychologist*; Ohlsson, S. (2009). Resubsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist, 44*(1), 20–40; Chi, M., & Brem, S. (2009). Contrasting Ohlsson's resubsumption theory with Chi's categorical shift theory. *Educational Psychologist, 44*(1), 58–63.

23. While there is considerable interest in concept change, the distinction between "to know" and "to believe" has a long history. People differ in the ways in which they distinguish knowledge from belief. One way in which this is done is in terms of degree of conviction. A discussion of this question far exceeds what we intend for this book. An interesting starting discussion can be found in Markham, A. B. (1999). *Knowledge representation.* Mahwah, NJ: Lawrence Erlbaum Associates (pp. 72–75).

Chapter 12
Encouraging Self-regulation

You probably have heard considerable talk about students' self-regulation in the context of learning strategies or teaching students how to study more effectively. Self-regulation is broader than this, however. It should be becoming clear from our portrayal of learning in the ULM that *at its heart, the ULM is about student self-regulation*. Students are the ones who are allocating their working memories and building their knowledge. Teachers and instructional designers can help students manage their working memory allocation and establish conditions that help direct attention and support motivation. Ultimately however, students manage and regulate their own engagement with the classroom, learning materials, and activities that we provide.

Five Students

What does student self-regulation mean in the classroom? If you are a teacher, think about the students you typically encounter. Rachel is a diligent "good" student. She works hard, uses effective study and learning strategies, learns things for the purpose of getting good grades and cares how well she does, but also wants to expand her own personal understanding. She is active in class, asks questions, seeks out help when she needs it, turns in completed assignments on time, and is never a behavior problem. All in all, she is what we would think of as one of our best students. She is probably taking honors or advanced placement courses. She might be our valedictorian. She is probably what we wish all of our students were like.

Sam seems to learn things and complete assignments easily. We may rarely see him actively studying, but he always seems to be reading or looking up things on the Web. He really wants to learn things; he will spend hours in the library or on-line following up on something that interests him. He examines everything in depth and likes challenges. He is active in class and asks questions. Sometimes Sam's are the best questions, ones that really dig into the material. He rarely seeks out or needs help. When he stays interested, he can do the best work in the class and get the best grades. But, when he gets bored, he slacks off completing assignments or turning in homework. This is most often true when he thinks something is just busy work. He

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_12,

might be as good as any student we have when he thinks there are opportunities to learn new things and grow, but he is a pain when things are moving too slowly for him. He might even be a class clown when bored.

Carol is a hard working student, one as diligent as Rachel. Carol doesn't seem to get the same results as Rachel, though. She studies hard, but not always effectively. Her studying focuses mostly on memorization for basic understanding. She rarely tries to get a deep understanding of the material. She doesn't learn things for herself. When she asks questions in class, they are mostly about what is going to be on the test or to find out what you, the teacher, think is important. She isn't really interested or personally engaged in the learning. Carol doesn't cause trouble and always seems to be working or studying, but she rarely stands out. Her work is adequate but rarely exceptional. She does well on tests, but she may not take much away from the course. She isn't really interested in doing extra things or supplemental activities. In many respects, we might not even know Carol very well. She is one of the large number of "average" students who do the work and don't cause trouble. We often don't pay much attention to them.

Dave just doesn't care about your class or maybe even school in general. He does as little as possible, just enough to get by. He doesn't seem interested or engaged by anything you are doing. He may or may not show up. He might be belligerent or disruptive, but more likely he just sits in the back not paying attention. You wish that you could figure out a way to connect with Dave, but nothing really seems to motivate him. His grades are marginal at best.

Finally, there is Bob. Bob always seems anxious. School is not a pleasant experience for him. He doesn't understand what you are teaching and doesn't seem to really know how to study and learn effectively. He looks to you and perhaps his fellow students for lots of help. He makes a lot of excuses and tries to get out of doing things. He doesn't seem very smart; he's afraid that he isn't smart. School is a very negative experience for Bob. He often is failing and seems lost much of the time. He lacks self-confidence and doesn't expect to do well. There is even a name for Bob in the literature: *learned helpless*.

There may be other kinds of students in your class, but we suspect that the majority of your students look like one or the other of these five archetypes. We have described them because research is beginning to identify these five *profiles* of self-regulation and the motivational factors that drive them as common to students in classrooms from elementary to post-secondary schools.[1,2]

Five Profiles of Student Motivation and Self-regulation

The detailed constellations of variables in the profiles of student motivated strategic self-regulation associated with the five students we described above are shown in Table 12.1.

These profiles reflect relatively stable patterns of motivation and self-regulation that students bring to the classroom. They might be thought of as a motivational and self-regulatory style; a preferred way that students approach a class. But while

Table 12.1 Profiles of students' motivated strategic self-regulation

Profile	Strategic	Knowledge Building	Surface Learning	Apathetic	Learned Helpless
	Rachel	Sam	Carol	Bob	Dave
Strategic self-regulatory behaviors					
General learning strategies	Yes	No	Yes	No	No
Deep learning strategies	Yes	Yes	No	No	No
Question asking	Low and high level	High level only	Low level only	No	Low level only
Help seeking	Adaptive	No	Adaptive	Expedient and avoidant	Expedient and avoidant
Lack of regulation	No	No	Yes	No	Yes
Study time and effort	High	Low	High	Low	Low
Motivation					
Performance goals	Approach	None	Avoid	Avoid	Avoid
Learning goals	Approach	Approach	Avoid	Avoid	Approach
Task goals	Approach	Approach	Approach	Avoid	Avoid
Expectancy of success	Medium	High	Low	Medium	Low
Means-ends expectancy	Effort	Interest	Effort	None	Luck; task difficulty
Self-efficacy	High	High	Low	Medium	Low
Outcome expectancy	Grades and learning	Learning only	Grades only	None	Grades and learning
Interest	High	High	Low	Low	Medium
Emotion/affect	Positive	Positive	Negative	None	Negative
Anxiety	Medium	Low	High	Low	High

these profiles can reflect relatively stable patterns, it is more proper to view them as situational rather than fixed traits or learning styles. Essentially all students possess the motivational and self-regulatory knowledge needed to assemble any of these profiles.

Students can change their profile over the course of a single class or in different classes and subject areas. One of the authors once worked with a college varsity athlete. This athlete practiced and worked out 5–6 hours each day. He was clearly following a strategic profile in relation to his efforts at further developing his athletic skills. As an academic student, however, he was apathetic at best. When asked if he thought that if he spent the same amount of time and put in the same amount of effort on his classroom work as he did on his athletic practice, he would be a better student, his answer was no. We all probably have encountered similar students. There are many students who are "Rachels" in all of their subjects except math, where they are "Daves." There is a possibly specious story about a college dean calling the chair of his physics department about a particular student who had been "spacing" his physical education requirement. The gist of the message was, "If you don't tell me Schwinger is likely to win a Nobel Prize, then he is out of here." The chair stuck up for the undergraduate. Issues connected with "... problems in completing the formalities" of that same student's PhD have been reported. Schwinger was a Sam/Bob who did go on to win the Nobel Prize.[3] It takes all kinds.

The motivational and self-regulatory knowledge in these profiles is retrieved like any other memory, through pattern matching to the particular sensory cues in the classroom environment. As the classroom environment changes across time or subject matter, we can expect different profiles to be retrieved. We have discussed previously in Chapters 4 and 5, how motivation and knowledge are constructed dynamically in working memory. Particular chunks of motivational knowledge including goals, means-ends expectancies, outcome expectancies, self-efficacy, interest, and so on are retrieved and lead to the construction of one or the other of the motivational patterns in the five profiles. This motivational construction then serves as a retrieval cue for the companion pattern of self-regulatory behaviors and effort.

These motivational and self-regulatory patterns are knowledge in long-term memories. As we acquire more stable probabilities in our expectancy, self-efficacy, and other motivational chunks, we develop more consistent, repeating, typical patterns of motivation for common classrooms settings, subject matter, and tasks. Self-regulatory behaviors such as learning strategies, question asking, and other forms of classroom engagement are procedural knowledge. These behaviors become linked to typical classroom and studying situations the same way any procedural knowledge chains are formed. For example, students use a learning strategy when reading to the extent that past use has resulted in desired learning or other outcomes. Typical motivational patterns drive use of self-regulatory strategies to the extent that specific patterns of strategies have led to means-ends expectancies about their effectiveness for achieving desired classroom goals, and self-efficacy in their effectiveness has developed. Sometimes the effectiveness of a strategy implementation needs to be explicitly discussed with students to be recognized as being successful. Students who have struggled to learn over a longer than expected period of time, many times do not attend to the fact that their strategy use was indeed successful. Past experiences have told them otherwise, especially when their strategy choices are inefficient for the particular learning task.

Over time, these typical motivational and self-regulatory patterns are retrieved as the "default" approach students take to classroom engagement and studying. These knowledge-based patterns interact in working memory with the situational environment of the classroom to produce the specific motivated self-regulation that a student is using at any particular time. If students have developed a *strategic* or *knowledge building* approach, like Rachel or Sam, these help sustain continuous motivation over time and keep working memory allocation directed toward appropriate learning goals in the face of difficulties encountered in the environment. Students who develop more problematic *apathetic* or *learned helpless* approaches, like Dave and Bob, place themselves at risk for failure.

Research into the development of these profiles is only beginning. Based upon what we know about how knowledge chunks and automated procedures develop, we can speculate that these profiles are likely to be more malleable in young children. As children grow older, we expect these patterns to become more stable. We also speculate that these patterns likely become more consistent across classes and subject matter. Even the profiles, themselves, may exhibit different degrees of stability and generality. Students who develop the *strategic* profile seem to follow this approach generally across all their classes. Good students seem to be good students in general. Other profiles like the *learned helpless* and *apathetic* may be confined more to specific classes or subject matter areas such as mathematics. These more dysfunctional profiles, however, can come to describe a student's approach to school in general; we certainly encounter "Daves" who nearly always seem apathetic about school, and "Bobs" who seem lost and failing in most of their classes.

What Can Teachers Do?

Since the self-regulatory behaviors in the profiles are fueled by motivation, the most direct impact that teachers can have is developing positive motivation, as we discussed in Chapter 8. To motivate students to engage in the more positive strategic and knowledge building self-regulatory profiles, a teacher can focus on getting students to

- establish learning goals
- develop a means-ends expectancy for effort
- develop outcome expectancies
- increase their interest
- embrace a positive affect

This is intended to move students toward one of the more positive strategic and knowledge building self-regulatory profiles. Fostering these motivational patterns can also help move students away from the more dysfunctional surface, apathetic, and learned helpless profiles.

Motivation itself is not enough. Teachers also have to help students develop effective self-regulatory skills, including appropriate metacognitive and cognitive learning strategies. A number of approaches to strategy training have been developed and should be consulted for specific techniques.[4] Teachers can also foster effective learning and problem solving strategies by verbalizing and describing their own thought processes when reading or problem solving. These can function as powerful examples and models for students to use. Overt modeling of thinking processes has been a key component of thinking skills teaching approaches such as reciprocal teaching.[5]

Beyond specific self-regulation and strategy training, aspects of the environment appear to influence which profiles students adopt.[6] Mastery goal structure in the classroom appears to facilitate adoption of a knowledge building profile. Performance goal structure appears to facilitate adoption of both positive and negative profiles as it increases adoption of the strategic, surface learning, and learned helpless profiles. Establishing a mastery goal structure appears to lead more unambiguously to a positive profile.

Student self-regulation can be shared through cooperative/collaborative learning.[7] Group members can share motivation and help motivate each other while compensating for individual deficiencies in specific learning skills or self-regulatory strategies. This can be especially important for young students who are still developing their personal self-regulatory skills. Increased student collaboration supports adoption of the strategic and knowledge-building profiles. Collaboration also supports the learned helpless profile, as these students use the group to cover their deficiencies in an unproductive manner. While collaboration and cooperative work can positively move students toward more productive motivation and self-regulation, it can also support dysfunctional self-regulation if students are not closely monitored and appropriate intervention taken when learned helplessness behaviors emerge. This type of dysfunctional behavior can be circumvented if the teacher designs a task incorporating individual and small group accountability, which is discussed just prior to the start of the collaborative activity. Specific roles within each group should be delineated to make clear what subtasks and responsibilities individual students will be held accountable. These tasks should contribute to the group effort for successful completion of the collaborative learning project.

Much forethought should, therefore, go into the planning phase of collaborative and cooperative work. Teachers should consider the range of abilities within each group to avoid making the spread too wide. Some variability is good, but too much of a range can often lead to inequities, with the high performing students doing much of the actual thinking while the low performing students are assigned to the artwork. It is wise to establish the parameters from the beginning for student participation, as well as to make available books, materials, electronic texts, and manipulatives with a wide range of reading levels.

Final Thoughts on Self-Regulation

Because research into these profiles is in its infancy, there is much that we don't know. This is especially true about what we can do in classrooms and through instruction to foster adoption of more positive profiles. We feel that these profiles provide coherency into the ways that the beliefs and emotions involved in motivation actually motivate certain patterns of behavior. The profiles offer a way to organize effectively the various motivation and self-regulation concepts and variables that are discussed in the literature. They also help illuminate the nuances in how differences in motivation lead to differences in strategic self-regulation. We cannot say for certain that these are the only profiles that may exist. These profiles do seem to reflect the ways most students behave. The same profiles have been found from elementary school to college. Therefore, teachers and instructional designers at all levels can use them to inform their attempts to help students be better self-regulated learners.

Notes

1. Shell, D. F., & Husman, J. (2008). Control, motivation, affect, and strategic self-regulation in the college classroom: A multidimensional phenomenon. *Journal of Educational Psychology, 100,* 443–459.
2. Moreno, R., Shell, D. F., & Pirritano, M. (2007, April). *Factors predictive of mathematics performance: Diversity between and within White, Hispanic, and Native American children*. Paper presented at the American Educational Research Association Annual Meeting, Chicago, IL
3. http://www-groups.dcs.st-and.ac.uk/˜history/Biographies/Schwinger.html (Accessed March 23, 2009).
4. Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Interventions with a focus on learning strategies. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 727–747). San Diego: Academic Press.
5. Hacker, D. J., & Tenet, A. (2002). Implementing reciprocal teaching in the classroom: Overcoming obstacles and making modifications. *Journal of Educational Psychology, 94*, 699–718.
6. Shell, D. F., & Husman, J. (2007, April). *Multilevel influences of students' dispositions and perceived classroom goal structure on their motivated strategic self-regulation*. Poster presented at the American Educational Research Association Annual Meeting, Chicago, IL.
7. Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64,* 1–35.

Chapter 13
Managing the Classroom Environment

In the same way that we learn reading and mathematics and science, we learn the expectations of a classroom – what we are supposed to do while we are there. When you walk into a classroom where learning levels are high, structure is apparent. Those classrooms may be quiet places or noisy places. There may be a teacher in front of a group, several groups being active, or all students working as individuals. When you scan around, however, *all* of the students appear to be engaged. Pressley and his colleagues provide a detailed description of what the productive hum looks like.[1] This does not happen automatically; management practices matter. *Because students learn management expectations, we expect the ULM principles and rules to apply to classroom management issues just as they apply to learning content.*

Structure – Based on Goals

In most classrooms, especially in primary grade classrooms, it helps when learning activities are structured such that once a student completes one task, he automatically knows to go on to another pending task without being asked to do so. Primary grade teachers should read Pressley's important book, *Motivating Primary Grade Students*, to familiarize themselves with ways in which this can take place.[2] In these situations, teachers establish learning goals that are quite open-ended as opposed to objectives that are limited and have a well-defined finality when successfully accomplished. Students expect to move from task to task without the expectation of being told to do so. Julianne Turner similarly found that when students have a sense of autonomy or choice and are offered optimally challenging work, they are highly motivated and engaged. They can explain what they are doing and what they are supposed to be learning.[3]

We have emphasized previously the importance of setting learning goals. But goals for task completion are also necessary if students are to get assignments and activities done in a timely manner. Our experience is that a little task structure often is helpful even for graduate students. Sometimes "bring me 60 pages between two pieces of cardboard by this date" is sufficient for getting a doctoral dissertation drafted. Sometimes "bring me a chapter summarizing the results without explaining them but including all of the data and statistical tests by . . . [a certain date]" works

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_13,

better. While over emphasis on tasks or even worse performance can be problematic if it devalues learning as a goal, school is a formal setting with deadlines and requirements. These need to be accomplished and teachers need to structure goals that motivate students to stay on task.

Contingency

The world can be a cruel place. We can do things that have every expectation of turning out well but that turn out poorly. A family is killed in a freak automobile accident while evacuating from a forecasted oncoming hurricane – that never arrives.

Generally speaking, however, life has contingencies. Students who study are the ones most likely to earn good grades in most school settings. College sports teams that practice generally outperform those that don't. Bridge pairs that play together often generally outperform those that are newly partnered.

Contingencies can be very tricky. Sometimes people succeed without preparation – for example, gaining admission to a college during a year when applications fall way off. Sometimes they fail in spite of preparation, such as not being admitted when there are many applicants for each spot. (How were you to know that the prestigious college you applied to attend would already have two freshman oboe players when you chose the oboe as your instrument?)

In your classroom, contingencies are the companions to goals. You need consequences for completing or failing to complete classroom activities and assignments. You mange these consequences by setting contingencies. That's why it is very important that the contingencies be understood. For example, if those late to class are never called upon, then some will be sure to be late when they are not prepared.

No matter how much we try to emphasize mastery, it remains a difficult issue. In many situations (the least favored subject in 5th grade, the required chemistry course in college, the required institutional faculty workshop on harassment), the students are there because some rule or person or requirement stipulated that they be there. For that reason, students are likely to have a task orientation rather than a learning orientation. Recently, one of the authors was in the campus student center getting a bite to eat before teaching an evening class, when a student and her peer math tutor sat at the table behind her. The student told her tutor, "I have to pass the test. I just want to know how to get the answers to these problems. You know, I'll never use this stuff again anyway." When task orientations like this student's prevail, the teacher needs to make contingencies for rewardable task completion both clear and absolute. If no papers are to be accepted after a specified date without some official excuse lest they incur a penalty, then be certain to penalize unexcused late papers. The best teachers nearly always pass some test of reasonableness; that is, there *are* some acceptable excuses. However, the looser this becomes, the weaker are the contingencies. Rest assured, students keep count – of what happens to them, and just as importantly of what happens (or they think happens) to their classmates. Sometimes an external person offers the advice, "Oh, they never do that," and the student is shocked to learn that the advice was wrong – or dated. *Always keep in* *mind: students count contingencies; teachers manage contingencies*. Consistency is your best bet.

Outcome Expectancy

Contingencies produce an expectancy of their continued future occurrence. Actions produce outcomes and we come to expect those outcomes. That outcome can be big or small. There was a conversation between two of our children about a favored, if some what quixotic, mathematics teacher who ultimately helped both children succeed. This teacher, JB, really hated it when students twirled their pencils, end over end. Those pencils inevitably would fall to the floor, and there would be a disruption. The teacher "went ballistic" when a student was caught twirling. One evening the younger child relayed how that had happened in her math class that day. The older child said, "Oh, he hates that. I remember one day when that happened. It was scary." Yes, as teachers you do become known for your outcomes, and especially when they are contingent upon certain well-defined student actions. There were two children who never twirled their pencils in math class, however automatic that seeming inane behavior had become, because the outcome was "too scary."

But another story about the same teacher shows how things work both ways. The younger child was remarkably good at mathematics – and often found herself bored in the classroom. One day she was awakened from a catnap during class to be asked about a problem on the board, and after a moment or two she blurted out "29." JB and her classmates laughed. Several minutes later, the problem evolved to $5^2 + 2^2$. Many of her classmates turned to her and stared. Based upon just this one incident, JB never called on her in the absence of her volunteering again. If the purpose of waking students up is to make sure they are on task and "getting it," no need to wake those who most likely are getting it. Better to say, "I'll let whoever wants to sleep sleep when they get it." Still better might be creating an environment where the sleeper is expected to study independently on some math-related topic or challenge project, or perhaps work as a tutor with another student having trouble or one who had missed class time due to illness.

Classroom Deportment

Managing classrooms is a complex and difficult task, one whose scope goes far beyond this book. As with so many issues, however, the ULM position is straight forward – help students *learn* the desired actions. We point you toward an extensive set of guidelines for classroom management that are consistent with the ULM.[4]

One of us recalls when our children were young, that the older child was in with a well-behaved group of students, but the younger child at the same school was in a group well described as being "wild." A new principal arrived and systematically undertook improving the deportment in that wild class.

Deportment has become a problem in college classrooms. Students reading newspapers and chatting always were a potential issue, but these situations pale in comparison with the impact of instant messaging on cell phones. Other issues discussed today are verbal or physical threats, to students or faculty, and incidents of disputing the instructor's authority or expertise. Today, most colleges have behavior codes for classroom deportment where such things were unknown several decades ago when a "no smoking" sign was the most code that was employed.

If you want to change a classroom behavior, you need clear consequences that are contingent on students' behavior. Consequences alone will not be enough, however. You need to get the students to think about that behavior. That is, for deportment changes as with nearly everything else school based, you need to engage working memory. Students need to attend to the issue. In many cases, they need to study the issue (repetition, feedback, etc.).

Teach Expected Behaviors

Teach expected classroom actions the same way as you teach anything else. If you were teaching US states and capitals, you would have posted a map of the states with each capital shown. If you were teaching chemistry, you would have posted a periodic table. If there are management or behavioral rules for your classroom, you should have them posted – especially in grades K-12.

We'll illustrate this with two examples. Suppose you want to call attention to all students in a first grade class. We've seen this managed in several ways. One teacher toggled light switches and remained silent – until all children were silent. Those who were not silent usually were quickly admonished (verbally) by those who otherwise were. Another teacher simply held up her hand with two fingers pointing up, and soon all children were holding up their hand with those same two fingers extended.

Let's turn to the other end of the spectrum, the college classroom. Plagiarism always has been an issue in college classrooms, beginning with copying homework and extending to other aspects of required assignments. With the advent of the WWW and the emergence of for-profit companies selling "term papers," plagiarism has increased. In fact, teachers can now participate in electronic services whose intent is to detect plagiarism.[5] Decades ago, students were expected to know what plagiarism was and that it was unacceptable. Today, nearly every college has formal statements about plagiarism.[6] Further, most schools expect teachers to include formal statements about plagiarism in their course syllabi.[7] If you assign learning activities where many references are brought together and quoted, then you should provide explicit statements about how you expect those citations to be made and include examples that you would consider to be plagiarism.

Dealing with Inappropriate Behavior

As a teacher, much of what was just written in this chapter did not come as news. In fact, you might even be saying to yourself right now, "I'll bet they've never met a kid like Xxxxx (fill in the name of your least favorite pupil).” Worse yet, some inappropriate behaviors have become automatic. That is, the students engaging in those actions really aren’t thinking about them.

Haven’t you ever been in a situation where you were thinking and then blurted out something you immediately regretted and wished you could take back? Sometimes we process automatically without considering how our outcome might be taken. Verbalizing your outcome may be automatic; once your thought is complete, out it comes. Yet we’ve all done this in a way best described as “putting our foot in our mouth.” Changing an inappropriate behavior that has become automatic is very difficult.

You would expect us to always say the same thing about poor behavior – get the student to think about it. If you want it changed, it needs to get into working memory. But it’s not at all simple. For example, it is clear that when dealing with one student in a classroom, you also end up dealing with all of the others as well. Whatever is done in public is done for all to learn from. What you do changes all other contingencies and all outcome expectancies. Maybe throwing a temper tantrum in a pre-calculus mathematics class is effective. Classes like that attract students who are there to learn the material. As already noted regarding JB, these brief tantrums can be both effective and memorable. In a different class with different students, maybe that approach would be downright foolish. Perhaps a student who is not succeeding in a different class might choose to twirl a pencil just to bring about an eruption – to control the teacher, so to speak. Worse, suppose the student who was getting everything decides to twirl a pencil instead of napping when bored and brings on an eruption; then what do you do? Following the guideline “Treat others the way that you want to be treated” is more effective than “Treat them the way that they treat you.”

Breaking Up Is Hard to Do

For those children for whom behavioral issues have become chronic, the ULM proposes to address the problem in a manner similar to that of experts engaging in deliberate practice. Once a behavior has become automatic, to change that behavior requires that the behavior somehow be brought into working memory making it susceptible to modification. When one examines the numerous cognitive-behavioral approaches to “behavior modification,” that’s precisely what they try to bring about. Essentially, each one of the approaches involves finding some way to pull an inappropriate behavior into working memory and think about it. Years ago, one of the authors taught in a preschool where the classroom had the “thinking chair” – a seat where students who made incorrect behavioral choices would sit to rethink and come up with more appropriate alternatives. They had to articulate the alternative and tell why it was a better choice. One colleague called it “the naughty chair.” This can be very difficult to do with bad classroom behaviors since they can play out without any conscious thought on the part of the perpetrator. In these cases, the strategies that seem to work involve tying the inappropriate behavior to a consequence that forces the behavior into working memory.

Perhaps the best example of behavior modification following the principles of the ULM comes from the Achievement Place Model[8] (AKA Family – Teacher Model; Boy's Town Model). The Achievement Place Model has been used successfully in a variety of settings including adolescent group homes and schools. It has been shown to be effective in almost 800 trials.[9] The Achievement Place Model uses a "token economy" where students are reinforced and punished by earning or losing points which are traded for tangible rewards or privileges to motivate behavior change. But Achievement Place anchors this motivation within a "teaching interaction" that draws on principles fully compatible with the ULM.

The teaching interaction is as follows:

Get the students' attention.
Remove points (to motivate sustaining attention to what follows).
Describe the problem behavior the student was doing (focus student attention on their behavior).
Have the student do the correct behavior (focused practice with guidance).
Reward the student by giving points (sustain motivation for practice).
Have the student repeat the behavior one or more times followed each time by giving points (motivate deliberate practice).

The teaching interaction is successful because it doesn't just punish for wrong behavior, it provides learning opportunities to acquire appropriate behaviors. We would argue that it is effective because it implements exactly the principles of the ULM.

While the desired outcomes of improving expert performance and changing an inappropriate behavior may at first seem poles apart, the reality is that much of each requires the same mental process – engagement in working memory – for the same reason; automated processes don't normally get into working memory where we can alter them. One of the things a psychological therapy often includes is maintaining a diary in which someone tries to record the circumstances and feelings of an untoward behavior. Perhaps, when a student has a problem area in which the teacher sees that problem coming on, having the teacher ask the student to make a diary entry could help. There would need to be some sort of sign for the teacher to make to initiate this, rather than just speaking to the student in a way that all present can hear.

We stress this one idea. As much as you may try to promise yourself that you will screen your speech for its social acceptability, an output – especially one that was automatic in coming – might just unleash itself inappropriately. The harder it is to come up with your response, the more likely the response is to "escape" because the less likely it will be that your working memory has held on to the "think before you speak" instruction. In accord with the ULM, *to modify an automatic behavior, the first thing we need to do is to somehow bring it into working memory*. If you are a habitual blurter who frequently suffers from foot-in-mouth disease, learning to control yourself will be difficult. The ULM explains why *it is much easier to learn a new habit than to break an old one.* It has to do with thinking about something rather than processing it automatically. When dealing with hard-core behavior that is inappropriate, getting the perpetrators to *think* about it at the time that their sensory inputs are most likely to initiate an automatic sequence that accompanies those inputs is the teacher's challenge.

Notes

1. Bogner, K., Raphael, L., & Pressley, M. (2002). How grade 1 teachers motivate literate activity by their students. *Scientific Studies of Reading, 6*(2), 135–165.
2. `Pressley, M., Kersey, S. E. D., Bogaert, L. R., Mohan, L., Roehrig, A. D., & Warzon, K. B. (2003). *Motivating primary-grade students*. New York: Guilford Press.
3. Turner, J. C. (1995). Turner,The influence of classroom contexts on young children's motivation for literacy. *Reading Research Quarterly, 30*, 410–441.
4. http://www.learningplace.com.au/uploads/documents/store/resources/res_35913_Modules_1-6.pdf (Accessed October 17, 2008).
5. http://www.safeassign.com/ (Accessed October 17, 2008).
6. http://www.unl.edu/gradstudies/current/plagiarism.shtml (Accessed October 17, 2008).
7. http://flwi.unl.edu/advice/avoidingplagarism.html (Accessed October 17, 2008).
8. Wolf, M. M., Kirigin, K. A., Fixsen, D. L., Blase, K. A., & Braukmann, C. J. (1995). The Teaching-Family model: A case study in data-based program development and refinement (and dragon wrestling). *Journal of Organizational Behavior Management*, *15*, 11–68.
9. Fixsen, D. L., Blase, K. A., Timbers, G. D., Wolf, M. M. (2001). In search of program implementation: 792 replications of the Teaching Family Model. In G. A. Bernfeld, D. P. Farrington, & A. W. Leschied (Eds.), *Offender rehabilitation in practice: Implementing and evaluating effective programs* (pp. 149–166). New York: John Wiley & Sons Ltd.

Chapter 14
Improving as a Teacher

Learning is about managing working memory. You can no more get into the heads of your students than you can get into your own head. Helping students manage their working memories as they learn, therefore, is a tricky task. Worse yet, if you have twenty students in your class, then you have twenty different working memories to help manage – each with its own prior knowledge, goals, chunks, and so forth.

Data on effective teachers is hard to come by.[1] While some teachers are standouts – both good and bad – it is hard to measure. In spite of this, it often is said that having one excellent teacher during the elementary years is enough to keep a student on a successful track for the rest of their careers. Highly successful students are those with learning goals who work to achieve those goals. The most successful teachers are those who get students to buy into learning goals such that, in the absence of teacher input, they continue striving to achieve the learning goal.[2] Conversely, having two really bad teachers in a row is an experience from which few recover. A friend once told this story about a flight back to California from a speaking engagement in New York City. To pass the time midway into the flight, she started chatting with the young man sitting next to her and found out that he was from a rough area in the Bronx. She noticed that he had been reading the Jewish bible. Because she thought this was an unlikely choice, she asked him what he was reading and why he chose to read it on this flight. He stated that he was going to start his freshman year at Stanford, and the book was required summer reading for the entire incoming freshman class. As they continued the conversation she found out that he was the first one in his family to even consider going to college. One question led to another and he said that, when he was in elementary school, he had a teacher who turned him on to reading. He'd become an avid reader since then. Such is the influence of one good teacher!

While important data regarding this issue is hard to come by, there is one consensus among researchers and the general public: *teachers matter*.

Despite the enigmatic nature of trying to determine what is "good teaching", the ULM is very clear about what good teaching is. To restate what we said in Chapter 7, all good teaching is anchored in the three principles of the ULM. Effective teaching facilitates and attempts to optimize students' ability to allocate their working memory capacity to the things we want them to learn. Effective teaching meets students

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_14,

at their current level of prior knowledge and engages students' knowledge productively for learning. Effective teaching supports and enhances students' motivation for learning and facilitates patterns of productive student motivated self-regulation.

The five rules of learning provide further guidance on what effect teaching entails. Effective teaching must do the following:

1. Teachers must help students focus attention on relevant materials and avoid distractions using the learning environment, instructional materials, and connection to students' prior knowledge. This includes the types of classroom learning goals, how instruction and materials are designed and presented, how classrooms are managed, and how teachers facilitate students' own self-regulation.
2. Teachers must provide opportunities for repetition of new information or learning processes. This includes providing multiple exposures to information through different instructional activities and discussion, providing for effective assessment, and giving appropriate feedback. For skills this includes opportunities for practice.
3. Teachers must provide ways for students to connect what they are learning to what they have previously learned in the class, what they have learned in other classes, and their other prior knowledge. Teachers need to help students construct meaningful connections between what they know and what they are learning. This includes providing opportunities for problem solving and thinking in the classroom.
4. Teachers need to provide support for maintaining students' motivation through the way information is designed and presented, use of engaging activities, opportunities for use of knowledge, and how feedback is delivered. Perhaps the single most important thing teachers can do to motivate students is to be themselves motivated by their subject matter and model that excitement and motivation to their students.
5. Teachers need to remember that learning is learning. Directing attention, providing repetition, facilitating connections, and providing motivation are what good teaching is about. Good teaching doesn't follow fads or take short cuts.

From the perspective of the ULM, teachers who do these things will be "good teachers" and their students will learn effectively.

Teachers' Prior Knowledge

Many, if not most, things that we learn involve help from teachers. Even with respect to such things as electronic games, experienced gamers often teach or give hints to novice players, sometimes through elaborate Websites. Coaches and mentors usually play vital roles for experts; the importance of teachers becomes evident through study of those at the very top of their fields, essentially all of who can identify their mentors and/or coaches.[3]

In teacher and school talk, *pedagogical content knowledge* is a description used often. Early work in establishing this label is attributed to Schulman.[4] It is sometimes said, "So-and-so really knows her stuff but, she just can't teach." What teachers need to get across a body of content is not just the content but also some know-how about teaching that content. The pedagogical content knowledge very much depends upon the setting: the pedagogical content knowledge required for teaching reading in pre-school is different from fifth grade mathematics or high school chemistry or graduate instructional design. Knowing how to teach reading in chemistry is even different from knowing how to teach reading in social studies! Consuming and producing texts well in one discipline is quite different from doing the same thing in another one. A P-12 pre-service teacher presumably can learn about appropriate ways to do this in methods courses.

Several issues emerge immediately. Teacher content knowledge is one. It is generally conceded that knowledge of fractions is a stumbling block for many students. Newton studied pre-service teachers' knowledge of fractions in five areas including computational skill, basic concepts, word problems, flexibility, and transfer. She concluded that teacher knowledge at the end of a pre-service course, especially with respect to flexibility and transfer, was "low."[5] Teacher mathematics content knowledge based upon degrees in mathematics and certification are positively associated with student learning in high school mathematics.[6] Primary grade teachers' mathematics knowledge is correlated with student learning outcomes.[7] Evidence for depth of content knowledge in areas other than mathematics is less strong. One possible explanation for the larger effects in mathematics is that this subject tends to be largely learned in school and usually has only small outside-of-school inputs.[8]

Methods Courses Versus Professional Meetings

The methods course context almost never exactly fits the ultimate learner. There's fifth grade mathematics and then there's fifth grade mathematics for an inner-city parochial school whose students are mostly Hmong immigrants. Even then, the students are individuals, each of whom will bring unique prior knowledge to the classroom. So, tips about one student's prior knowledge are unlikely to apply to all students.

Still another issue about methods course approaches is that they rarely take into account prior knowledge and motivation in a systematic way. So, fifth grade math teachers may have some activities that seem to work for their students, and chemistry teachers know of classroom chemical demonstrations that seem to create situational interest (i.e., are motivating).

Applying the lens of the ULM to much of what is considered to be pedagogical content knowledge taught in methods courses will lead to revisions of that content.[9]

An effective place for teachers to acquire pedagogical content knowledge seems to be at professional meetings attended by experienced teachers with very similar teaching assignments. The specificity shared at these meetings seems to lead to pedagogical content knowledge transfer. Successful teachers become more successful by learning from one another.

Video Clubs

Because all teachers do not have ready access to professional meetings, other modes of professional development should be sought out to increase the levels of pedagogical content knowledge across content areas. Attendance at professional meeting can present particular problems for teachers in rural schools where funds and substitute teachers are not plentiful. One practical option for almost any size school is to form "video clubs" within a school or district.[10] Teachers within these groups use videos of their teaching to stimulate discussions about teaching and learning in the content areas. These discussions can be further enriched and informed through common readings from journals published by content area professional organizations and books based on current research in teaching.

Attention: New Learning Requires Working Memory

You do not always have to work on new learning; it *can* just happen. That's what episodic memory (autobiographical memory) is all about. While incidental learning *does* happen, it's really inefficient. Moreover, in the absence of some remarkably dramatic event, episodic memory is both fragile and susceptible to modification.

If you want something to "stick," it needs to be in semantic memory, and the transition from episodic to semantic memory requires effort. You must allocate working memory to it. We attend to the top slot in working memory. To learn something you need to pay attention to it; it's got to make its way into working memory and become the focus of attention while there. For example, the role of attentional control in reading comprehension recently has been addressed.[11]

What kinds of things focus attention? For those students seeking grades, just saying "This will be on the test," usually is enough to garner attention. The drawback to this is in terms of focusing on performance (test scores matter) rather than learning. One of us recalls an event from 50 years ago when a roommate who was pre-med returned from a biology class during which the teacher had spent an entire period discussing his research.[12] After making several unflattering comments about the lecture, he scrawled a single word in large letters over each of eight carefully written single line pages of handwritten lecture notes: *Forget*. While this talk probably worked in terms of providing situational interest for a few students in the large lecture class, it certainly did not work for all, and not for that roommate.

An attention-getter may need to do two things: capture attention by relating new material to prior knowledge, and do it in a way that encourages later recall. The best way to find attention getters is to talk about prior experience. Sometimes a local happening or event in someone's life provides a source of attention. Generally speaking, the best way to find attention getters is to discover from other successful teachers what works. We've already mentioned unusual zoo animals as examples.

Repetition

Repetition is required for long-term memory. If something is going to go into semantic memory, repetition is required. Repetition must also include retrieval and an attention to the spacing of repetition.[13] Research in vocabulary instruction has shown that learners need about seven encounters with a word to place it firmly in their memory and be able to retrieve it. Recall that Nuthall and Alton-Lee determined that students needed at least four repetitions of information in varied contexts for it to be learned.[14] Clearly, repeated exposure to the word seven times in one lesson is not as effective as exposures over a few days and weeks. A somewhat contemptuous description that has emerged in writing about school is "drill and kill." The drill part refers to repetition. The kill part usually implies "kill interest." Even today some societies consider learned persons to be those who have memorized enormous amounts of content. When no connections are important – as usually happens with rote memorization – the accepted strategy is rehearse, rehearse, rehearse, repeat, repeat, repeat, drill, drill, drill. It works! Heaven knows, it's not a lot of fun – at least until after you have accomplished a difficult goal such as being able to recite *the Gettysburg Address* from memory and can show that off.

Unless the particular learning goal *is* rote memorization (as with *The Pledge of Allegiance*, *The Lord's Prayer*, etc.), a successful teacher tries to find ways for students to repeat the same information that go beyond rote memorization. Earlier we cited the work of Beck and McKeown from their *Text Talk* project that begins with contextualizing a word by talking about its use in the story. ("In the story you will hear about how the farmer trudged home after a long day of work.")[15] In that example, the word used (trudge) was repeated in several different contexts, and each repetition also afforded new opportunities for making connections.

Sometimes you really do need drill. Whenever some knowledge must be automated, then it will need to be practiced over and over. Think about Tiger Woods practicing his putting – over and over and over. When an engineer is performing some calculation, he really can't afford to allocate resources to decide what five times seven is. Much expert knowledge *must* be automated. Also, recall the dramatic result from retesting all Swahili words, not just those that the learner was unsure of.[16]

Again, by talking with other teachers you'll find ways in which repetition has been included in games and other activities. It's never a one-size-fits-all proposition except when rote memorization is involved.

Connections

Learning is about connections. That's something we consider crucial in the ULM. Sometimes making connections is rather easy. For example, it is easy to see how you would connect the notion of angles with the notion of triangles. Even the names speak to the connections. And, of course, students know about connections. For example, Fred knows that if he throws a spitball at Sally's hair, John will break up laughing and the class soon will be in an uproar. We all learn about connections, big and small.

Sometimes we just fail to encourage making connections. For example, in college chemistry laboratory experiments where students are expected to discover, say, the "molar mass of oxygen," the students often gather data and then often leave to write up a report. They never discuss the experiment or the results; they never are asked overtly to connect the things they have just done in lab to other things they have learned or are learning.

Sometimes we fail to connect the knowledge we have learned to those who discovered that knowledge. Most people have heard about Newton and Einstein. Few people have heard about Eric Kandel, the physician-scientist whose work provided much of the basis of our understanding of how neurons work and without which this book could not have been written.[17]

Connections to content often come in textbooks. We mention Kandel; to learn about his work, check the reference and read his book. Imagine what it must be like to sit in a Kandel seminar and hear about the ins and outs of his discoveries. Today, Web searching has made the task of making connections quite a bit easier. Students can be assigned the task of searching out connections.

Content connections are usually varied but often included in complete curricula. They can be fun.

Effortless Versus Requiring Effort

The ULM asks that you think about your teaching in new ways. Of those new ways, *perhaps none is more important than the notion of effortless versus requiring effort*. Effortless knowledge, knowledge that comes from our autobiographical memory system, is fragile and susceptible to alteration. It usually is not the goal of classroom instruction. Although semantic memory requires effort, it usually is the kind of knowledge we seek to impart in school. Not only does the content have to be brought into working memory and attended to, but this has to happen more than once. It's harder still if that knowledge needs to be automated.

There are many activities that students look forward to. The first draft of this section was written on Halloween. At that time, the chemistry education listserv was filled with discussions about a chemical demonstration.[18] One takes potassium chlorate, a somewhat dangerous chemical, and heats it in a test tube until it melts. Then one carefully drops a sugar candy – often a Gummi bear – into the melt. The reaction is explosive.[19] When you eat a Gummi bear, you also make carbon dioxide – but you do it in a far less dramatic way. We have the classic case for situational interest. It's Halloween; kids know about Halloween candy and trick-or-treat. We humans oxidize eaten candy, but just not in quite this spectacular way. There are big flashes, loud noises; both fumes and excitement fill the air. Assuming this is done safely (it is *very dangerous* when not done properly), a good time will be had by all. At the dinner table, the student might say to her parents "Ms. Jones blew up a Gummi bear in chemistry today." If something goes wrong, it may become the talk of the school or even the local news.

Here's the issue. In the absence of any other instruction, a month later the best students will have only a vague recall of the details of what transpired. They may remember the day for the rest of their lives, but the details of what they recall are likely to be far removed from the reality of what was done. Is this a problem? Well, it depends. If the goal was just to have a good time, then it is not a problem. If the goal is to better understand oxidation reactions, then this activity needs much more follow through. The facts need to be repeated. The explanations need to be repeated. There might even be a student question or two that needs an appropriate answer. Perhaps the students need to write a report, or write a sentence or two about this on a test or in some homework.

School can be fun; it need not be burdensome. When you are a teacher, you need to keep in mind that there is a balance between episodic memory (after all, it's *all* episodic when it comes in) and semantic memory where we are trying to store information for later use. One goal you should have is to tie activities together so that they involve repetition and connections.

Connections Are Connections

Learning is learning. In the sense that we are speaking about them, connections are connections. We take the view that too much is made of innate ability. Further, the literature does not really support the notions of either multiple intelligences or learning styles. At the same time, it is clear that some people are (or at least think they are) visual learners while others learn better from hearing and so on. We assert that, if this is the case, it is because of what they have learned rather than what they were born with.

Our "take" on modalities is different. You *do* have different sensory inputs. Sight, sound, smell, and touch *are* different. Each has inputs; each has "buffers;" each can come into working memory using one slot. You can "think" about each.

In Chapter 5, we described an experiment where students study and were tested in the presence of odors (chocolate, camphor).[20] Camphor-taught students tested better in camphor, and chocolate-taught students tested better in chocolate. How could this happen when odor had nothing to do with the content studied? As before, the ULM explains this observation in terms of pattern matching that is available for recall at testing. Having the same odor present gives just one more thing for a pattern to match against, and so goes the explanation. If you don't share a neurobiological view of learning as espoused in the ULM, then it is very difficult, if not impossible, to understand the result of this and similar experiments. Pavio made a reputation developing the dual coding theory.[21] Our view is that we should think of an n-coding theory in which all senses may play a role. Information from all senses can be stored, "played back," and included in pattern matching.

The reason to teach using different sensory modalities and test using different modalities is not really rooted in intrinsic differences among us humans; learning is learning. What does matter, however, is that all of us benefit when the connections we make involve as many inputs as possible. Connections in multiple modalities make learning easier.

This also makes practical sense. Much of what we do in schools ends up being measured by pencil-filled spots on papers of scanable response sheets. Virtually no one makes a living filling out such sheets. Instead, we apply whatever we have learned by writing or painting or talking or wielding a scalpel or whatever. Connections in multiple modalities make learning more realistic and potentially more useful. Internists must be familiar with issues of male prostate disease. In addition, they need to be able to palpate a prostate as part of a screening to detect such disease.

Closing Thoughts

When you are seeking ways to enhance your pedagogical content knowledge, try to frame what you are learning according to the ULM. For example, how will this new content comport with making connections? Does it afford you a good way to offer repetition? Will it generate the kind of situational interest likely to bring it into focus in working memory? If a suggestion looks like a good one for episodic memory, how can you follow it up with something aimed at moving it into semantic memory?

Ultimately, human learning works the same way for all of us. Yes, there are differences in working memory capacity, prior knowledge, and motivation. But, whatever a learner's current knowledge state, the three principles and five rules apply to new learning.

Notes

1. For example, see paper available here: http://www.tqsource.org/link.php (Accessed March 23, 2009).
2. Pressley, M., Kersey, S. E. D., Bogaert, L. R., Mohan, L., Roehrig, A. D., & Warzon, K. B. (2003). *Motivating primary-grade students*. New York: Guilford Press.
3. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*. New York: Ballantine Books.
4. Shulman, L. (1986). Those who understand: A conception of teacher knowledge. In B. Moon & A. S. Mayes (Eds.), *Teaching and learning in the secondary school* (p. 386). New York: Routledge.
5. Newton, K. J. (2008). An extensive analysis of preservice elementary teachers' knowledge of fractions. *American Educational Research Journal, 45*(4), 1111–1154.
6. Monk, D. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*, 125–125.
7. Hill, H., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

8. Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237.
9. Gess-Newsome, J., & Lederman, N. (2001). *Examining pedagogical content knowledge: The construct and its implications for science education*. New York: Kluwer Academic Publishers.
10. Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education: An International Journal of Research and Studies*; *24*(2), 417–436; van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' 'learning to notice' in the context of a video club. *Teaching and Teacher Education: An International Journal of Research and Studies*, *24*(2), 244–276; Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education: An International Journal of Research and Studies*, *20*(2), 163–183.
11. Conners, F. A. (2009). Attentional control and the simple view of reading. *Reading and Writing*, *22*(5), 591–613.
12. *Zygocotyle lunata*, known to college pre-meds as Willey's worm.
13. McGaugh, J. L. (1966). Time-dependent processes in memory storage. *Science 153*(3724): 1351–1358.
14. Nuthall, G., & Alton-Lee, A. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal*, *32*(*1*), 185.
15. Beck, I., & McKeown, M. (2006). Different ways for different goals, but keep your eye on the higher verbal goals. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 182–204). New York: Guilford Press.
16. Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968.
17. Kandel, E. R. (2006). *In search of memory: The emergence of a new science of mind*. New York: Norton.
18. For an interesting discussion about using chemistry demonstrations, see Ramette, R. (1980). Exocharmic reactions. *Journal of Chemical Education, 57*, 68–69.
19. http://www.youtube.com/watch?v=CJ-pSfXcXtw&NR=1 (Accessed March 23, 2009).
20. Schab, F. R. (1990). Odors and the remembrance of things past. *Journal of Experimental Psychology, 16*(4), 648–655.
21. Pavio, A. (1990). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Chapter 15
Policy

You might think that it would be a bold, giant leap to translate the parameters that the ULM ultimately places on learning into establishing policy. Actually, this is far from the case. Most of the logical extensions of the ULM into policy have been recommended, sometimes for years, and some have been studied.

Pre-school

Prior knowledge is the biggest factor in the ULM. Ability counts, but most people have enough core working memory capacity to succeed at nearly anything. Motivation counts, but is influenced strongly by one's prior knowledge and experience. Any policy that is likely to enhance prior knowledge should have payoffs. Preschool is an obvious policy. From the perspective of the ULM, preschool helps build a basis of literacy and general knowledge that can help expand working memory capacity in elementary school. Preschool helps to level differences in prior knowledge, especially in the areas of foundational vocabulary and concept development.[1] Preschool can also help develop positive motivations toward learning. Finally, preschool can help establish basic behavior routines. Students can transition smoothly to formal school settings with fewer distractions and less extraneous load.

One of the best-known studies was conducted in Michigan with "at risk" children in which the study paired participants with similar non-participants. (A discount rate takes into account the cost of money, and is a way of helping to judge investments):

> . . .The data show strong advantages for the treatment group in terms of higher lifetime earnings and lower criminal activity. For the general public, gains in tax revenues, lower expenditures on criminal justice, lower victim costs, and lower welfare payments easily outweigh program costs. At a 3% discount rate the program repays $12.90 for every $1 invested from the perspective of the general public; with a 7% discount rate, the repayment per dollar is $5.67. Returns are even higher if the total benefits—both public and private—are counted. . . . [2]

These results need considerable explication. A very large chunk of the economic benefit comes from reduced expenses due to incarceration; as a consequence there is significant gender difference.

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_15,

The Partnership for America's Economic Success estimates that the ultimate payback for investments in preschool is compounded at such a significant rate that preschool is at the cornerstone of this group's policy recommendations.[3] Nobel Laureate James Heckman has written a review that brings together much of what is known about the impacts of health and early schooling.[4]

Schools

Modern schools in the United States are about many things, of which instruction and learning are just two. Whatever the reasons for this may be, at least part of the justification for this is rooted in the fact that instruction is regarded more as an art than a science. If you buy the arguments made in this book, then much obfuscation about learning goes away. Learning involves managing working memory. Teaching involves helping learners to manage their working memories. How can we translate such an abstract notion of a model into meaningful policy?

School: Discourage Words Suggesting Innate Abilities

The last principle of the ULM is "working memory allocation is directed by motivation." As we have indicated, motivation depends very much on our beliefs. Everything we can do to support the notion that effort pays is likely to lead to enhanced learning. Put differently, there won't be anything more than episodic memory-based learning without effort. We work in a college where one of the most advocated slogans is "build talent into careers." The policy should change and move away from notions of "talent" and "smart" to those of effort. Talent inevitably implies something from birth:

talent:[5]
(a) special – often athletic, creative, or artistic aptitude
(b) general – intelligence or mental power, natural aptitude or skill

As we have previously discussed, one notion that becomes quite clear from detailed studies of experts and expertise is that talent really means skill, and that skills come about as the result of hard work and effort more than something we possess at birth. We've pointed out several references that essentially question the importance of innate talent.[6] Ross' *Scientific American* article provides an especially readable discussion of the development of talent.[7]

Children need to know and believe that learning is something they control through their effort. They need feedback and reinforcement that they can succeed, even if the going gets tough. In the ULM, the statement "all children can succeed" is not just a slogan; it is a fundamental principle. A good school policy, therefore, would be to *avoid using words like talent, smart, and IQ* when discussing student learning. One might say that they place the em*PHA*sis on the wrong syl*LAB*le.[8]

District, State: Organize Based upon Knowledge

The second principle of the ULM is that working memory's capacity for allocation is affected by prior knowledge. It is clear that a child arriving as a five-year-old at kindergarten who reads and comes from a rich literacy environment and one who arrives there with minimal reading from an environment with no books at all are not likely to make the same gains. This likely has nothing to do with ability and little to do with motivation. That is, both children can come from environments highly supportive of the notion that school is where you go to learn how to read. From a purely scientific viewpoint, expecting the same interventions for both children to achieve the same final outcome is unrealistic. It won't happen! It has nothing to do with teachers or desires; it has to do with how humans learn. When we understand learning from the perspective of the ULM, we can better clarify expectations for schools and project more reasonable learning outcomes for children. Preschool has come to be recognized for very large economic and social impacts.[9] Why? Preschool helps to level differences in prior knowledge. School year differences in rates of learning seem small when compared with differences that emerge over summers, especially with respect to reading.[10]

Currently schools are organized around calendar age. In the United States students start school in kindergarten when they are about five years of age, and they proceed through 12 grades. The physical organization of those schools varies, but usually includes elementary school, middle school (or junior high school), and high school. Some savvy parents tend to hold back those of their children who might just make it into kindergarten for an extra year before they start. Indeed, this sometimes makes a big difference in a child's apparent success. Gladwell makes a point of how many professional Canadian hockey players are born in the early months of the year (January–March) and how this increases the likelihood that they will become selected for development. One of us had this sort of experience with the game T-ball[11] where kids hit balls propped upon stands and then run bases akin to the game baseball. One year "our" team lost every game but went on the next year to win every game. As it happened, the kids during the second year were the same as they were in the first; the team roster was nearly identical. So, the entire team was older, more developed, and more experienced.

When children are sorted into groups in the current age-based structure, it more often than not is done on the basis of ability than anything else – so-called ability grouping. Ability is typically determined by some type of "intelligence" test. Ability grouping is controversial and has a history of mixed success.[12] There are a few situations (cross-grade reading, mathematics) where success seems significant, but across-the-board grouping (i.e., all subject areas) does not appear to be beneficial.

Rather than either age or intelligence/ability, it makes more sense from the ULM's learning perspective to group children according to what they know, and to let them flow through this grouping based upon both what they know and come to know. To make this work there would need to be several "out points," where a child could be described as "having finished" a curriculum. This used to be done in a quite different way, where students sought different high school diplomas. Until recently, New York State was the only state with content-based exit requirements in the early 1990s.[13] Indeed, New York graduates did demonstrate higher levels of content knowledge than those of other states. They also generally fared better economically.

Any system of schooling essentially works to the advantage of those students coming from homes that nurture and support learning. Conversely, schools do not deal well with students from unsupportive home environments.

A system that allowed differentiated time periods to achieve learning goals could work to the advantage of those otherwise disadvantaged learners. In essence, longer times, extra time in school, and the attendant costs would minimize socioeconomic differences. Allowing students to flow through based upon learning also would accommodate differences in prior knowledge that arise from differences in advantage. The reality of the differences between current school practices and what the ULM tells us about learning is this: schools focus too much on what they call ability meaning intelligence or some other entity type talent and not enough on effort. As we stress in the ULM, in the end, ability really is prior knowledge, what one knows, and that comes about through effort.

Over the last few decades, special advantages become apparent during summers when school is not in session. Advantaged children can attend special programs for athletics, music, debate, science, computing, drama, and an enormous range of similar activities. These serve to magnify differences in knowledge based upon social or economic status.

The notion of organizing schools along the lines of what learners know certainly is not new. For example, B. S. Bloom recommended it in 1966:

> . . . but it is evident that the sequence of learning tasks from grade to grade is not given the same kind of attention that is now given to the sequence of learning tasks within a grade. This undoubtedly is an organizational problem arising from the assembly line notion of education in the graded system of education. Sequence in learning is not just the avoiding of unnecessary repetition or overlap from grade to grade. It is the planned movement of learning from one level of complexity or mastery to another. The development of sequence in learning requires not only the planning of subject matter and materials over time; it also requires the development of continuity in teacher-student relationships over time. . .[14]

While this type of individualized approach to school has been problematic to do in practice, technology may provide ways to implement a truly individualized, self-paced curriculum in the context of the formal school setting. We are likely not there yet, but it is possible to envision a virtual school where classes are composed of students drawn together from multiple physical locations, perhaps from around the world, based on similar levels of prior knowledge. These free floating classes would continuously recompose as students left or entered based on progress acquiring the subject matter knowledge.

Gifted Students, Skipping Grades, and Advanced Placement

Many school systems offer gifted programs; these vary in many ways. A report on such students sets forth its position in the title, *A Nation Deceived: How Schools*

Hold Back America's Brightest Students.[15] This two-volume document reviews the current national status with respect to gifted education. As you would expect, the ULM supports anything done for these students so long as a few guidelines are followed. First, the gifted label should be awarded based upon achievement with the regular curriculum rather than some test of ability or intelligence. Next, if the curriculum is a solid one, then the student should have an opportunity to move through *that* curriculum. Finally, if resources are available for individual students, payoffs are more likely to come from investing in the low performers rather than the high performers. That is, invest in those for whom mastery is a challenge.

The report, one with substantial research support, splits out 20 "most important" points. The first three are:

1. Acceleration is the most effective curriculum intervention for gifted children.
2. For bright students, acceleration has long-term beneficial effects, both academically and socially.
3. Acceleration is a virtually cost-free intervention.

One way to begin to identify those students likely to benefit from acceleration is to look at achievement scores and think about what is best for those students at the very top.[16] Rather than use an ability or IQ measure to select those for acceleration, however, it makes more sense to use a related-content test for older students. The report provides a discussion about this issue. The economics of this are rather clear. It probably costs $250–$500 to decide that a student should be accelerated. If the student skips an entire year, that is one less year the student will be in the district. Assuming that much of this one-year matriculation cost is recoverable, it can be applied for those with problems achieving mastery. Districts probably could have more of these resources if they paid the accelerating student. The risk here, of course, is in engendering performance rather than learning goals by having the students working for money instead of learning.

Twenty-five years ago, we offered placement tests in general chemistry. We used a standardized test and insisted on something like two thirds of the items being correct. We administered the test prior to having students pay the official fee to take the test. We risked abuse, but the test had a reputation such that no one with a poor chance of passing took the test. If the student passed the test, they then would pay their fee and be able to transfer the tuition to some other course (second semester chemistry, first semester physics, etc.). On balance that system worked very well for students, faculty, and the university – because most of those skipping "by examination" were really good students and most chose to stay for a full program.

The *Deceived* report advocates *Advanced Placement*. *Advanced Placement* programs have become increasingly popular. Because we are in a mid-sized city, many local advanced students prefer to take university courses. The young woman mentioned earlier for her math skills graduated high school, having completed college calculus and one semester each of differential equations and linear algebra. This gave her academic advantages she never lost. Though students like her were rare, they were by no means unique; there were five to ten who made similar accomplishments that graduated at the same time she did.

The downside of *Advanced Placement* programs is that they can overly focus on performance and task goals rather than learning. Too many *AP* courses epitomize "teaching to the test" consisting of little more than drill and practice for the *AP* Exam. This is far different from an accelerated or advanced curriculum based on more in-depth learning. Little in contemporary Advanced Placement programs seems directed at developing a "life long desire" for learning the subject matter. How much of the subject matter in *Advanced Placement* is remembered after the *AP* test is over and the college credit is received? We suspect little.

From the perspective of the ULM, a "gifted" program should be about learning more and in more depth. It should challenge students to acquire greater depth and breadth of knowledge and facilitate creation of larger more complex knowledge structures in memory. It shouldn't be just about getting through faster. But if students are able to learn in depth more quickly, the ULM would support opportunities for them to accelerate their studies.

State, National: Large Pre- and Primary School Impacts

Any system of schooling essentially always works to the advantage of those students coming from homes that nurture and support learning. Conversely, schools do not deal well with students from unsupportive home environments. Oftentimes, this ends up tracking socio-economic status. For example, SAT scores correlate with mean family income.[17] As a result of this, one way to assess the impact of any intervention is to express it in terms of gain (or loss) in equivalent family income. That is, intervention xyz has the mean effect of raising the family income by, say, $20,000.

Using such a yardstick, preschool, interventions with respect to mathematics have been shown to have a powerful impact.[18]

> Previous work with this sample had shown that the effect of 1 year of part-time preschool was equivalent to increasing family income by more than £10,000 (£ = British Pound Sterling) a year.[19] We show that the effect of primary school was even more important than preschool (0.39 versus 0.26 SD), but both were sufficiently large to be important for any government wishing to maximize educational achievement. They are greater than the effect for father's education and similar to that for family income but less than that for mother's education (see figure, page 1161). Analyses for low and higher income groups reveal that the effects for the HLE and preschool and school effectiveness are remarkably similar for both income groups, which indicates their importance across the income spectrum. These effects are predictive, but we cannot assume causality. Observational studies, such as this study, do not have random assignment, so it is always possible that results may reflect selection bias and/or the operation of unmeasured variables.[20]

Early-Career Teacher Mentoring

A teacher may end up having a 30–45 year teaching career. Some students must have that teacher during the teacher's first year or two of teaching. Generally, those years are not so good. While there is a theory to pedagogical content knowledge, practice is quite different; we never know what to expect.

After decades of teaching experience, one of us was with our Harvard-professor daughter when she was awaiting papers to be turned in at the end of first semester. She had two sections of eight students each, and was expecting a total of 48 papers. Two days before her deadline, only ten or so papers had been turned in. We made a small $20 wager in which “the experienced one” said that at least 10 papers would not come in, thinking that would be a “sure thing.” All 48 papers were turned in on time. A sure thing at one college may be a sure loser at another. Successful strategies vary; expectations vary.

Moreover, when you begin, you’re just not very good compared to what you are likely to become. As we noted before, it takes 10–15 years to become expert. This applies to teaching just like any other profession. If you work at your teaching, you’ll become better at it, just as you become better in the kitchen or better on the golf course.

Every school would like to have a policy of never hiring new teachers but only those “broken in” elsewhere. Where we live, this happens – with large, urban schools often raiding rural schools for teachers once they’ve had some experience. In large urban areas, the reverse is often true. Large inner city schools often get the new teachers as the experience ones flee to the suburbs, which of course accounts for part of the problems inner city schools have.

While teachers of sophomore organic chemistry often complained about what was going on in freshman general chemistry – just one year of teaching freshmen almost always had them return to sophomores with a better appreciation of both sophomores and freshman chemistry instructors.

To make the system work better, policy should support mentoring of teachers early in their careers. Teachers identify many resources as being important during their early years, including mentors. Mentoring probably should extend to experienced teachers during their first year in a new setting. This of course implies that all schools have a mix of experienced and new teachers, with experienced teachers in the majority. This may also be the reasoning behind why one East Coast private school – The Benchmark School – requires their new teachers to apprentice with a master teacher for up to two years, then co-teach with a master teacher for one additional year before being designated a master teacher and assigned a classroom of their own.[21]

Policy Summary

The three core principles of the ULM are:

1. Learning is a product of working memory allocation.
2. Working memory’s capacity for allocation is affected by prior knowledge.
3. Working memory allocation is directed by motivation.

When policies are developed for schools, they need to take these into account. Schools are places where students allocate working memory. That is, they are places where students pay attention to things society expects them to know about: reading, arithmetic, appropriate behavior, democracy, and so forth. True enough, states in the United States have their own curricula, but they end up being similar to one another in most ways.

Schools are places where students acquire knowledge. A child's environment outside of school before and during the school years has a tremendous impact on what knowledge that child is likely to acquire in school. Policy needs to take that into account.

Perhaps most important, schools and school boards like "smart" kids. But, the science behind the ULM shows that effort matters more – much more. And the "smart" kids most often are smart because of what they know – the results of their prior efforts at learning, not because they possess some special ability. Policy needs to take that into account – right down to the ways in which schools and governments speak about learning.

Notes

1. Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Brookes Publishing Company.
2. Nores, M., Belfield, C., Barnett, W., & Schweinhart, L. (2005). Updating the economic impacts of the High/Scope Perry preschool program. *Educational Evaluation and Policy Analysis, 27*(3), 245.
3. http://www.partnershipforsuccess.org/index.php?id=01 (Accessed March 23, 2009).
4. Heckman, J. (2007). Economics of health and mortality special feature: The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences, 104*(33), 13250.
5. http://www.merriam-webster.com/dictionary/talent (Accessed March 23, 2009).
6. Bloom, B., & Sosniak, L. (1985). *Developing talent in young people*: New York: Ballantine Books; Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown, and Company; Ericsson, K., Prietula, M., & Cokely, E. (2007). The making of an expert. *Harvard Business Review, 85*(7–8), 114–121; Colvin, G. (2008). *Talent is overrated: What really separates world-class performers from everybody else*. New York: Penguin Group.
7. Ross, P. E. (July, 2006). The expert mind. *Scientific American, 295*, 64–71.
8. At the very end of the writing process, one of us noted that teachers often take tests intended to measure certain "talents" prior to hiring. We have addressed repeatedly the subject of talent. It is our view that data support an incremental model for talent rather than an entity view; that is, "talent" as we use this word in everyday language is something that can change through learning. The "other view" is expressed by those who advocate instruments to measure talent: "Skills can be learned, and knowledge can be obtained. However, talent – the key to strength and peak performance – must exist naturally within a person. A talent is a naturally recurring pattern of thought, feeling, or behavior that can be productively applied. They are spontaneous, top-of-mind, perhaps even subconscious, reactions to situations. Talents are what one does well 'without even thinking about it.' They are innate, non-teachable."[From http://www.careertrainer.com/Request.jsp?lView=View Article &Article=OID%3A113426 (Accessed March 23, 2009)]. The *Gallup Teacher Perceiver* Interview was often used by school districts, sometimes with respect to hiring practices. A meta-analysis of a version of the instrument in use until a few years ago was studied, and a "modest relationship (r=0.28) between the TPI and some measure of teaching quality" was reported. (Metzger, S. A., & Wu, M.-J. (2008). Commercial teacher selection instruments: The validity of selecting teachers through beliefs, attitudes, and values. *Review of Educational Research*,

78(*4*), 921–940.) Metzger and Wu state: "This could mean that the TPI chiefly reflects a teacher's ability to be liked by an administrator." [p. 931]. We have appealed to the scientific literature in making a case to reject the entity notion in favor of an incremental notion. It is widely acknowledged that a purpose of education is to develop skills in many areas, both those in which the student already has strength as well as those in which strength still is lacking. "Gallup research has proven that the best way to develop people – and net the greatest return on investment – is to identify the ways in which they most naturally think, feel, and behave as unique individuals, then build upon those talents to create strength, the ability to provide consistent, near-perfect performance in a specific task." [http://www.gallup.com/consulting/61/Strengths-Development.aspx (Accessed March 23, 2009)]. When leading an organization, one's goal more likely is to optimize what is at hand rather than to enhance employee skills in diverse areas. Thus, the kind of advice a manager is likely to obtain from Gallup is understandable in the context of what one might hope to learn from a management consultant. The *preponderant* United States view of talent is an entity attribute – you have it or you don't. In many ways, the TPI supports this view. There is an irony in this. The districts that employ tests such as the TPI most often do so in an attempt to identify prospective teachers that might have pro-student attitudes. *This pro-student view implies an incremental notion of ability*.

9. Nores, M., Belfield, C., Barnett, W., & Schweinhart, L. (2005). Updating the economic impacts of the High/Scope Perry preschool program. *Educational Evaluation and Policy Analysis, 27*(3), 245.
10. Alexander, K., Entwisle, D., & Olson, L. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 23*(2), 171–191.
11. http://www.teeballusa.org/What_is_TBall.asp (Accessed March 23, 2009).
12. Slavin, R. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research, 57*(3), 293.
13. Bishop, J., Moriarty, J., & Mane, F. (2000). Diplomas for learning, not seat time: The impacts of New York Regents examinations. *Economics of Education Review, 19*(4), 333–349.
14. Bloom, B. (1966). Stability and change in human characteristics: implications for school reorganization. *Educational Administration Quarterly, 2*(1), 35.
15. Colangelo, N., Assouline, S. G., & Gross, M. U. M. (2004). A nation deceived: How schools hold back America's brightest students. http://www.accelerationinstitute.org/Nation_Deceived/ND_v1.pdf; http://www.accelerationinstitute.org/Nation_Deceived/ND_v2.pdf (Accessed March 23, 2009).
16. Paula Olszewski-Kubilius, "Talent searches and accelerated programming for gifted students," Chapter 7 of Volume 2 of *A Nation Deceived*.
17. Jaschik, S. (2007). SAT scores down again, wealth up again. *insidehighered.com*. http://www.insidehighered.com/news/2007/08/29/sat (Accessed March 23, 2009)
18. Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M., et al. (2008). The early years. Preschool influences on mathematics achievement. *Science, 321*(5893), 1161–1162.
19. Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). *The effective provision of pre-school education (EPPE) project: The final report*. London: Department for Education and Skills (DfES).
20. Supporting on-line material: http://www.sciencemag.org/cgi/content/full/321/5893/1161/DC1 (Accessed March 23, 2009).
21. Gaskins, I. W. (2004). Professional development at Benchmark School. In D. S. Strickland & M. L. Kamil (Eds.), *Improving reading achievement through professional development* (pp. 195–213). Norwood, MA: Christopher-Gordon.

Chapter 16
Frontiers

A model such as the ULM opens up new ways of thinking about the kinds of activities that have been going on in classrooms, schools, training centers, museums, homes and nearly everywhere – since learning is an everyday happening. Here we let ourselves wander far afield from our original purposes.

Declarative Versus Procedural

A couple of us have gone around and around on the seeming distinction between procedural and declarative memory. Rather than go down this path, perhaps we can try the following: Some memories result from sensory inputs and/or processing. That is, a person sees or hears something, or she thinks about it and stores the result of the thinking. Someone else, however, can only know that the person has stored this by asking him to do something. That means as teachers we only know students have learned something when they can "output" this to us by telling us, writing it down, demonstrating it, doing a behavior, or what-have-you. This implies that declarative memory/knowledge is not especially meaningful until it can be connected to some type of procedural output. We have discussed this in Chapter 4 as proceduralization that involves paring declarative knowledge to motor or cognitive outputs. Procedural memory involves outputs and declarative memory can only be output through a procedure. As a result, the end point of any instruction, even that involving only declarative learning of facts or concepts, is the establishment of appropriate procedures for expression of the learned knowledge. As we delve into the declarative and procedural distinction at the neurological level, it would seem that declarative memory can exist entirely in the cerebrum but that procedural memory probably always has a neural path that includes the cerebellum which coordinates motor outputs.

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_16,

Where Is Working Memory?

We have been vague about where working memory is located. When working memory works, that tissue always is connected to something else that also is working, such as a sensory input or some stored prior knowledge. When experiments are performed to detect working memory while it is working, multiple regions are involved. Over a decade ago, one region started to become implicated systematically in working memory function. The frontopolar cortex, located in the frontal lobes behind the top of the forehead, is a region of great interest.[1] This region is implicated in analogical reasoning[2] and subgoal processing.[3] Some workers recently claimed: "The frontopolar cortex (FPC), the most anterior part of the frontal lobes, forms the apex of the executive system underlying decision-making."[4] There are reports trying to connect IQ with brain volume in which the volume of this area is positively correlated with IQ.[5]

Parental Involvement; Rote Memorization

To many people, memory means what cognitivists think of as rote memory. That is, the term memory often is thought about in a very narrow way. We need to worry about this terminology, since we use the term in a very different and far broader way. It probably is worth thinking about rote memorization.

Rote memory has one learning strategy – rehearse, rehearse, rehearse. It's an easy strategy to learn even if it is very difficult to enact. For example, it is alleged that many young Muslims memorize the Koran (77,801 words).[6] It is asserted that this process takes youngsters 2–3 years. Similarly, Asian immigrant and Asian American students generally outperform traditional classmates, an outcome ascribed to greater parental involvement and emphasis on education. In these cultures, there is great emphasis on rote learning. Rote learning counts in many quarters, however. Recall the "learning success" described for the child who learned how to recite the names of the U. S. Presidents.

When parents who themselves are not especially well-educated set educational goals for their children and become involved in the schooling process, it is likely that there is an emphasis (overemphasis) on rote memorization. The process is reasonably understood; involvement means checking out the learner's rehearsal.

Laboratory Teaching

There are many issues in contemporary K-12 education in the United States (and elsewhere) that might benefit from a view through the ULM lens. The U. S. National Academy of Sciences released a report in 2006 concerning the status of K-12 laboratory teaching.[7] This report drew several conclusions including:

> Conclusion 2: Four principles of instructional design can help laboratory experiences achieve their intended learning goals if: (1) they are designed with clear learning outcomes in mind, (2) they are thoughtfully sequenced into the flow of classroom science instruction, (3) they are designed to integrate learning of science content with learning about the processes of science, and (4) they incorporate ongoing student reflection and discussion.

In the lens of the ULM, having clear outcomes (1) fits with the notion of having appropriate goals. Ongoing reflection and discussion (4) really means thinking about the work, and tying that new learning to prior learning. Thoughtful sequencing (2) amounts to appropriately accommodating prior knowledge. Integrating science content and process (3) amounts to the outcome of reflection and discussion. At a surface level, there is no difference between the ULM and this conclusion. One could have started with the ULM and come up with this.

We see a serious complicating factor, however. The fourth ULM rule is that learning is either effortless or requires effort. Effortless learning is that coming into episodic (autobiographical) memory. The nature of classroom laboratory activity is such that it is rich with views of oneself. More than three decades ago, we found remarkable differences in the way students viewed themselves in lecture and laboratory settings. Students use first person pronouns (I, me, we) in their descriptions of laboratory work much, much more than in their descriptions of lecture.[8] At the time we were interested in explaining anomalies in the teaching evaluation process. It never dawned on us to consider this as a study in the difference between episodic and semantic memory, or, expressed in the terms of the fourth learning rule, effortless learning versus learning requiring effort.

As already noted, everything learned starts off as episodic memory.[9] Furthermore, when it comes right down to it, there are no processes of science as such. Instead, there are processes of microscale synthetic organic chemistry, processes of forensic crime scene investigation, etc. Let's assume that we can somehow arrive at suitable content for teaching laboratories. Early on, it might be good to have learners decide, "I can learn how to do this work," and later, "I can do this work." In other words, since laboratory work already seems to be so inherently autobiographical, perhaps establishing explicit goals in the "I can learn/I can" category is appropriate. These could help boost self-efficacy as they focus students on the process of learning prior to focusing on the product.

The Science Demonstration

An often-used science classroom activity is the demonstration (lecture demonstration). In such an activity, the lecturer performs or manipulates materials or devices, and speaks of some principal of science. In chemistry, liquids turn colors or explosions occur or light is emitted – with whatever the phenomenon being demonstrated almost always being aptly described as "novel." Ramette went so far as to publish several papers describing "exocharmic" reactions, those that seemingly exuded charm.[10] Chemistry teachers can develop a significant following based upon their skills as demonstrators. Hubert Alyea was often regarded as the highpoint of the class reunions at Princeton, for example, based on the alumni's recollections of his demonstrations and the promise of an interesting show.[11]

While reports of demonstrations abound, and large science departments often employ a full-time professional "demonstrator," studies of learning from demonstrations are few and far between. Roadruck's discussion of learning from chemical demonstrations includes the following exaltation:

> The demonstration, in so far as it is performed by the teacher and not the student, is at best a vicarious experience for the student. A demonstration cannot simply be presented. To achieve its purposes in developing an understanding of the content, a demonstration must be part of a lesson that gets students actively involved with the event and the content it seeks to impart. The demonstrator must involve the students in the experience through questioning, predicting, redesigning, the offering of explanations and the testing of those explanations.[12]

This paragraph calls for what the ULM considers to be conscious allocation of working memory resources. Here's the problem: Because of the novelty, and especially with very exocharmic demonstrations, episodic memory is almost certainly going to store parts of the learner's experience during the demonstration. The bigger question about instruction relates to what this will be stored with. Successful storage of the conceptual material with that seen and/or heard and/or smelled from the chemical demonstration is likely to require significant processing, repetition, connections with other material, and so forth.

Because episodic memory is quite subject to distortion, it is not surprising that aspects of demonstrations remembered long into one's life past the initial event are, in fact, heavily modified – and some would say, incorrect. It is not at all surprising, then, that chemistry teachers are chagrined when they hear first a former student's recollection of some classroom demonstration, and then have that followed by a seriously erroneous description of the demonstration.

The moral of the story for teachers is that when something that is effortless (like a demonstration) is paired with something that requires effort (such as the underpinning conceptual explanation), you need to pay more attention to that which requires effort – if teaching for semantic memory is your goal.

Informal Education

What about learning outside of school? We learn from television and radio, newspapers, and from just living life. One place that many societies pay attention to is the museum. According to The International Council of Museums:

> A museum is a non-profit, permanent institution in the service of society and its development, open to the public, which acquires, conserves, researches, communicates and exhibits the tangible and intangible heritage of humanity and its environment for the purposes of education, study and enjoyment.[13]

What might the ULM say about informal education in a place such as a museum? We look at this from the vantage point of classroom learning. How do museums fit into what we expect from classroom learning?

First, except for the scholars who come to utilize collections for advanced study, most museum learning is effortless. That is, it depends upon episodic rather than semantic learning. So, as with similar settings, one would expect an outcome from museum exhibits to involve changes in motivation. That is, one goal outcome ought to be a change in affect toward some aspect of the collection.

Another way to view museums is to consider them as starting points for further study. In that case, we should expect museums to ensure that some sort of follow-up to a visit is possible, one that supports the development of deeper learning.

These ought to be the kinds of performance measures that museums gather. That is, change in attitude and follow-up learning based upon visits.

Evaluating Teachers

The movement to evaluate K-12 teachers and reward them based upon performance is gaining momentum. College teachers have been evaluated for decades, and these evaluations have had some impact on the college faculty reward structure. When a student goes to school, many outcomes are expected. One of these is learning. That is, how do we evaluate teachers based upon learning?

College Teacher Evaluation

College teachers are evaluated mostly if not exclusively on the basis of student evaluations. Generally speaking, these are well accepted. However, student evaluations clearly depend on student episodic memory. College learning goals, generally, are based on semantic memory. A system of evaluation of learning that makes sense probably should include measures of learning gains. This is a tricky business. Students who are not motivated toward learning or performance are unlikely to learn. Episodic memory probably is a good way to measure student motivation through self-reporting. On the other hand, there are situations where students' feelings are positive but the amount of or quality of material learned is not very great. Based upon the ULM, attempts at measuring learning would be an improvement if included in the teacher evaluation process.

K-12 Teacher Evaluation

We've referred in this book about the "common wisdom" about spontaneous learning in autistic savants, a folklore wisdom that almost stopped our efforts. There is another folklore wisdom that most of these authors hold about teachers. If a child has one really good teacher in primary school, that child is set for life. One year is long enough to embed and sustain habits of mind that can serve one for a lifetime. On the other hand, having two really bad teachers back to back causes a loss from which essentially no child recovers. We concede that the data upon which we base these judgments is highly flawed.[14]

The last decade has seen tremendous emphasis on test performance and achieving high test scores. The good news is that, for the most part, the performances are not normed but based on achievement. The bad news is that the scores measure end-of-term knowledge rather than learning gains. Generally speaking, students in those schools located in upper-middle socioeconomic status neighborhoods outperform those from less advantaged areas. This often is translated to reflect good teaching. Is this fair? We've already noted the work of Alexander et al. who studied semester gains in learning where the differences in summer gains between students from different SES groups were dramatic.[15] Again, the ULM accounts for these differences in terms of prior learning. Gladwell would account for these in terms of differences in opportunity and legacy.[16] The point is that teachers have little to do with the out-of-school differences, and it is unfair to hold them accountable for these.

On the other hand, anyone who has visited schools and watched a variety of teachers work will tell you that there are vast differences between classrooms in terms of engagement, quantity, and quality of student work. We do advocate maintaining measurements of learning that inform us about how students are achieving relative to absolute learning standards. For teacher evaluation, however, we advocate measuring learning gains against population-specific expectations of learning gains. Do not take this as advocacy for warehousing low-performing students. As we indicated in our policy positions, to the extent that the mission of schools is to impart students knowledge, that is what we need to be measuring. We can deal with this either by offering different levels of graduation achievement (differently labeled degrees), accept the fact that some students will take longer than others to accomplish designated learning levels, or some combination of these two. While we never can have a system that flawlessly provides and measures learning, we can do much better than we now do in terms of focusing on, accomplishing, evaluating, and rewarding learning.

Quantitative Modeling

Unlike many educational learning "theories" that have been put forth, because the ULM is based on mechanisms that have been identified at the neural level and on computational cognitive models similar to Newell's *SOAR* mentioned in Chapter 6, the ULM ultimately could be developed into a fully realized unified theory of cognition. While this task is beyond the scope of this book, it remains a future endeavor.

Notes

1. Buckner, R., Kelley, W., & Petersen, S. (1999). Frontal cortex contributes to human memory formation. *Nature Neuroscience, 2*, 311–314.
2. Green, A., Fugelsang, J., Kraemer, D., Shamosh, N., & Dunbar, K. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research, 1096*(1), 125–137.
3. Braver, T., & Bongiolatti, S. (2002). The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage, 15*(3), 523–536.
4. Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*, 594–598.
5. Haier, R., Jung, R., Yeo, R., Head, K., & Alkire, M. (2004). Structural brain variation and general intelligence. *Neuroimage, 23*(1), 425–433.
6. Luo, M. (2006). Memorizing the way to heaven, verse by verse *The New York Times*. http://www.nytimes.com/2006/08/16/nyregion/16koran.html (Accessed October 2, 2008).
7. Singer, S., Hilton, M., & Schweingruber, H. (2006). *America's lab report: Investigations in high school science*. Washington, D.C.: National Academy Press.
8. Levenson, H., & Brooks, D. W. (1975). Student evaluation of lecturers *versus* graduate laboratory instructors in introductory college chemistry. *Journal of College Science Teaching, 5*, 85.
9. There are some aphasics who have no autobiographical memory but who can, after many repeated trials, learn. See: Reinvang, I. (1985). *Aphasia and brain organization*. New York: Plenum Pub Corp.
10. Ramette, R. (2007). Exocharmic reactions up close. *Journal of Chemical Education, 84*, 1; Ramette, R. (1980). Exocharmic reactions. *Journal of Chemical Education, 57*, 68–69.
11. http://tigernet.princeton.edu/˜ptoniana/alyeaobit.html (Accessed March 23, 2009).
12. Roadruck, M. (1993). Chemical demonstrations: Learning theories suggest caution. *Journal of Chemical Education, 70*(12), 1025–1028. See also O′Brien, T. (1991). The science and art of science demonstrations. *Journal of Chemical Education, 68*(11), 933–936.
13. http://icom.museum/definition.html (Accessed March 23, 2009).
14. For data that support our assertions but are of doubtful quality, see: Hancock, K. (1998). Good teaching matters... a lot. http://www2.edtrust.org/NR/rdonlyres/ 0279CB4F-B729–4260-AB6E-359FD3C374A7/0/k16_summer98.pdf (Accessed March 23, 2009). For a discussion of teacher evaluation, see: Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washing, D. C.: National Comprehensive Center for Teacher Quality http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf. (Accessed December 20, 2008).
15. Alexander, K., Entwisle, D., & Olson, L. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 23*(2), 171–191.
16. Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown, and Company.

Chapter 17
Epilogue

We opened our discussion of the ULM with the question: What is learning? In this book, we have presented an answer: a theoretical model of learning that we call "the Unified Learning Model" (ULM). Learning in the ULM is a relatively permanent change in a person's knowledge. This knowledge at its most basic biological level is the relatively permanent change in a neuron or group of neurons. Understanding that learning is the change in a neuron complex in the brain allows us to understand that learning is a product of those things that change neurons. These are attention, repetition, and connection. Attention, repetition, and connection occur primarily in working memory. Thus, the first principle of the ULM anchors learning in working memory; specifically how limited working memory capacity is being allocated. Working memory capacity however is not static. It is changed by knowledge. As the second principle of the ULM notes, the more knowledge we have, the more working memory capacity we have. Learning begets more learning and begets more complex knowledge as the capacity of working memory expands to allow integration of more and more knowledge in larger interconnected chunks and networks. Finally, according to the third principle of the ULM, we allocate our potential working memory capacity according to how we are motivated.

The principles of the ULM define what can be thought of as *mental contiguity*. This notion of the ULM as mental contiguity takes a statement such as "hands-on; minds-on" from a platitude to something potentially meaningful. The ULM defines what it means to have your "mind" on something.

When we say that a student needs to allocate their working memory to the learning task, what we mean, in one sense, is that students need to be "in contact" or contiguity with what we want them to learn. Contiguity is the basic principle in both the classical conditioning of Pavlov and the operant conditioning of Skinner, as well as fundamental to the association rule of working memory.[1]

Contiguity is also behind some of the more prevalent findings from the "process-product" research in the 1960s and 1970s on the relation between teacher behaviors and student outcomes. These studies found that time on task and amount of curriculum covered during the school year were the major determinants of student outcomes.[2] Students can't learn something if they never come in contact (contiguity) with it, whether in school or elsewhere.

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7_17,

What the ULM does is refine the notion of contiguity from a purely physical presence (the student is in the same place as what is to be learned) to a *mental* presence (the student needs to have their "mind" in the same place as what is to be learned).

Essentially, students who do not engage with their learning according to the principles of the ULM might as well not be there. They may be in the classroom or performing an activity, but if their minds are not attending to information, doing repetitions, and making connections, they might just as well be somewhere else. Physical contiguity is necessary in school learning. That is, the student has to show up in school for anything to happen. That is not sufficient, however. Learning only happens when students have mental contiguity through the principles of the ULM. It is this mental contiguity that we are trying to motivate students to engage in. It is this mental contiguity that teachers are trying to facilitate through their instructional methods, materials, media, and activities.

The ULM espouses three principles of learning:

1. Learning is a product of working memory allocation.
2. Working memory's capacity for allocation is affected by prior knowledge.
3. Working memory allocation is directed by motivation.

These three principles are enacted through five rules of learning that provide specific guidance for realizing them in teaching:

1. Direct student attention to the desired knowledge to be learned. Help students focus attention on relevant materials and avoid distractions through the learning environment, instructional materials, and connection to students' prior knowledge.
2. Provide necessary repetition. Provide multiple exposures to the knowledge to be learned and opportunities for recall and practice.
3. Facilitate connections. Provide ways for students to connect what they are learning to what they have previously learned in the class, what they have learned in other classes, and their other prior knowledge. Help them to construct meaningful connections between what they know and what they are learning.
4. Provide motivation to facilitate effort. Recognize that learning can be difficult and provide support for maintaining students' motivation.
5. Remember that learning is learning. Directing attention, providing repetition, facilitating connections, and providing motivation are the parts of good instruction. There are no short cuts; good teaching and instruction do not follow fads.

Teachers and other educators who follow the ULM can ensure that their students have their "minds on" their learning, whether they have their hands on or not. This mental contiguity will maximize student engagement and ultimately student learning and achievement.

Notes

1. King, D. B., Viney, W., & Woody, W. D. (2009). *A history of psychology* (4th edn.). Upper Saddle River, NJ: Pearson Education; Leahey, T. H. (2004). *A history of psychology, a: Main currents in psychological thought* (6th edn.). Upper Saddle River, NJ: Prentice Hall.
2. See Good, T. L., Biddle, B. J., & Brophy, J. E. (1975). *Teachers make a difference*. Oxford, England: Holt, Rinehart & Winston.

Index

D.F. Shell et al., *The Unified Learning Model*, DOI 10.1007/978-90-481-3215-7,