Science of Learning, Retention, & Sustainment Training

Table of Contents
Science of Learning .................................................................................................................. 3
Overview ...................................................................................................................................... 3
Topics in Learning Science ......................................................................................................... 4
  Constructivism ............................................................................................................................ 4
  Models of Memory ....................................................................................................................... 4
  Situated Cognition ....................................................................................................................... 5
  Deeper Learning and Learning Outcomes ............................................................................... 6
Learning Environment Design .................................................................................................... 8
Assessing for Understanding ...................................................................................................... 8
Adaptive Expertise ....................................................................................................................... 9
Personalization and Individualization ....................................................................................... 10
Cognitive and Intelligent Tutoring ............................................................................................. 14
Transfer of Learning .................................................................................................................. 16
Learning Science Research Methods .......................................................................................... 17
Example Learning Science Related Funding and Work .............................................................. 17
  Examples of work funded by NSF or the Department of Education ........................................ 18
Learning Retention .................................................................................................................... 19
  Interventions that Support Better Retention ...................................................................... 21
    Problem-based, authentic, situated, and active learning ..................................................... 21
    Teaching for conceptual understanding and deep conceptual knowledge ..................... 22
    Repeated retrieval of information ....................................................................................... 22
    Performance aids .................................................................................................................. 22
Impact of Learning Science on Sustainment of Skills and Knowledge ...................................... 23
  Training Transfer Considerations ......................................................................................... 24
  Evaluation of Training ............................................................................................................ 25
  Assessment and Cost Analysis in Training ............................................................................. 26
Interventions to Improve Sustainment of Skills and Knowledge ......................................................... 27
Teaching for Deep Conceptual Knowledge .............................................................................................. 27
Performance Support or Aiding .............................................................................................................. 28
Use of Cognitive Tutors .......................................................................................................................... 28
Big Data and Learning Analytics .......................................................................................................... 28
Conclusion ............................................................................................................................................... 30
References ................................................................................................................................................ 30
Science of Learning

Overview

Beginning in the 1960’s, research from psychology, computer science, philosophy, sociology, and other scientific fields began to converge around learning and education. As they closely studied learning, researchers began to discover that instructivism or the current modes of thought and practice concerning education and learning were (in their opinion) deeply flawed. In the 1990’s the following ideas about learning from this convergence reached a consensus and were published by the U.S. National Research Council:

- The importance of deeper conceptual understanding (only acquiring facts and procedures is not enough)
- Focusing on learning in addition to teaching (active participation in the learning process)
- Creating learning environments (key features of design for deep conceptual understanding)
- The importance of building on a learner’s prior knowledge (prior knowledge must be engaged or misconceptions persist)
- The importance of reflection (developing knowledge should be expressed, reflected and self-analyzed) (Bransford, Brown, & Rodney Cocking, 2000; Sawer, 2006a).

This convergence has expanded to include the fields of cognitive science, neuroscience, and educational technology and instructional technology. Learning science was considered “born” in 1991 when the first international conference was held and with the first publication of the Journal of the Learning Sciences (Sawer, 2006b). Learning science (LS) is an interdisciplinary field that focuses on furthering the empirical understanding of learning or the science of learning. In general, the goal for learning science is to better understand the cognitive and social processes that result in the most effective learning and then using this knowledge to redesign more effective learning environments (Sawer, 2006b). The original research foci of LS include cognitive-psychological and social-psychological foundations of human learning, as well as on the design of more effective learning environments. Over the years, these have expanded to include the design of curricula, informal learning environments, instructional methods, policy innovations, digital environments, and machine learning. Contributing fields to learning science include epistemology, cognitive science, artificial intelligence, computer science, educational psychology, neuroscience, and anthropology. There are various approaches and theories about what learning is, how it occurs, and how it can be validated. Taken from the various disciplines, these approaches and theories tend towards those concerned with the human mind (psychology, epistemology), modeling of the mind (cognitive science, artificial intelligence), the human brain (neuroscience) and biological systems, human social systems (sociology, anthropology), and human technology intersections (educational technology, instructional technology, human computer interaction). This report discusses topics in LS, examples of LS-related funding and work, learning retention, interventions that support retention, impact of LS on sustainment of skills and knowledge, assessment and cost analysis in training, interventions to improve sustainment of skills and knowledge, and conclusions.
Topics in Learning Science
Learning science (LS) has essential foundations and ways of thinking. These foundations are all situated within a constructivist paradigm (Kuhn, 1996) both epistemologically and methodologically. Typically learning scientists work within their area of interest, but the community is collaborative and cross-pollination is prevalent.

Prominent foundational areas described in *The Cambridge Handbook of the Learning Sciences* are implicit learning and the brain, informal learning, and designs for formal learning and beyond. The integration of these areas is seen as a key to developing a transformative theory of learning (Bransford et al., 2006) and leans heavily on insights from neuroscience and social theories of learning. Other foundational areas described in *The Cambridge Handbook of the Learning Sciences* (Sawer, 2006b) are cognitive apprenticeship, cognitive tutoring, and learning in activity.

Constructivism
Constructivism is described as a theory of knowledge rooted in philosophy, psychology, and cybernetics (von Glasersfeld, 1995). Constructivism also has roots in cognitive science, neuroscience, and other disciplines. As a theory of learning and knowledge, it mirrors the makeup of the LS community, guides its activity, and defines its research methods. Constructivism operates on the assumptions that knowledge is actively constructed cyclically from experience and new knowledge is either assimilated or accommodated into these constructions. In the words of Ulric Neisser, "The central assertion is that seeing, hearing, and remembering are all acts of construction, which may make more or less use of stimulus information depending on circumstances." (1967, p.10). Meaning is constructed through interactions with the world including (and importantly) social interactions. It is considered rooted in, and indexed by, experience. Learners in this sense are not empty vessels to be filled but active organisms constructing meaning (Driscoll, 2000). However, meaning is imposed by the individual on the world rather than existing independently in the world (Duffy & Jonassen, 1992). A constructivist viewpoint sees the individual constantly adapting to the situational demands of the world. This point is illustrated by constructivists using the concept of autopoiesis described by Maturana and Varela (1987) and Winograd and Flores (1986): the structural coupling of an organism with its environment for successful existence.

Models of Memory
Constructivist models of memory are conceptually different from traditional cognitive information processing (CIP) theory. Somewhat analogous to a computer, in CIP information is accumulated through sensory input into visual or auditory sensory memory. Through attention and pattern recognition, this information is moved to short-term or temporary working memory. Through processes such as rehearsal and chunking, information is then encoded in long-term memory for later retrieval. Long-term memories are based on network models of representation or semantic models (Driscoll, 2000) which are commonly referred to as ontologies.

Constructivist models of memory are based on connections and experience. Much like a rhizome, this model of memory allows for multiple, non-hierarchical entry and exit points for representation and interpretation. Roger Shank argued that through juxtaposition in memory, any two things may be linked with meaningful relationships generated between them (Shank, 1988 in Driscoll, 2000). For a more
tractable perspective, a particular slice of the “rhizome” would reveal a person’s knowledge at that time in that context and there would be no assumptions of memory invariability over time or across contexts. However, this viewpoint presumes that neither knowledge nor its descriptions are stable. It suggests that students do not learn rules but connections which resemble rule-based performance to satisfy constraints of experience and environment.

Other models come from cognitive science and, like CIP, are based on what we know about computer memory techniques. A central feature is the idea of a pointer or a way for one memory location to point to or refer to another location. This technique leads to the nesting of hierarchical data structures with the highest level structure containing pointers to simpler, lower level structures thus providing a metaphor and example for how knowledge might be modularized from the viewpoint of cognitive scientists (Sawer, 2006a). This metaphor may support Shank’s idea of juxtaposition.

As these ideas began to take hold, coincidentally John Anderson, a founder of a line of research on cognitive tutoring at Carnegie Mellon University, evolved his theory of knowledge and cognitive architecture called Adaptive Control of Thought-Rational commonly known as ACT-R, which is compatible with the assumptions of cognition and memory in constructivism (Budiu, 2013; Driscoll, 2000). Also the work of Shank and Kolodner in scripts and cases is based on constructivist notions of memory discussed in Kolodner’s work in case-based reasoning as a theory of knowledge (Kolodner, 2006) and as a method for implementing intelligent systems.

**Situated Cognition**

From the perspective of LS and constructivism, knowledge is situated in activity occurring within social, cultural and physical contexts. Therefore learning, especially deep conceptual learning, occurs through activities and tools situated in these contexts. To learn history, a learner must “do” history using the tools and methods of historians. These activities are referred to as authentic practices (Sawer, 2006a). If the learning goals are to successfully troubleshoot large IT systems, then the learner should actually engage with the activity of troubleshooting authentic simulated versions of a system and, ideally, progress through several systems learning the basics of troubleshooting across systems. However, the concept of situated cognition is much broader than these examples. It incorporates learning communities and methods such as anchored instruction and is related to social-cultural psychology, activity theory, distributed cognition, and ecological psychology. For researchers focusing on learning in activity, the primary unit of analysis is an activity system: a complex social organization containing learners, teachers, curriculum materials, software, tools, and the physical or virtual environment (Greeno, 2006). However, situated cognition is a fundamental concept guiding the aforementioned foundations of LS.

Evidence of learning gains from a LS perspective can be found in learning activities requiring the construction of knowledge using disciplined inquiry to produce discourse, products, or performances that have value beyond school. Chicago students who received assignments requiring authentic intellectual work made greater than average gains in reading and mathematics on the Iowa Tests of Basic skills (ITBS) and in reading, mathematics, and writing on the Illinois Goals Assessment Program (IGAP) (Newmann, Bryk, and Nagaoka, 2001 in Brookhart, 2010).
Problem solving, reasoning, critical thinking, and the active and reflective use of knowledge are seen as the goals of constructivist designs for learning. To reach these goals several conditions of learning or learning principles should be present. They include the following:

- Complex, realistic, and relevant environments
- Social negotiation as an integral part of learning
- Multiple perspectives and the use of multiple modes of representation
- Ownership in learning (student centered, relevant)
- Self-awareness of the knowledge construction process (metacognition)

A constructivist approach can be contrasted with the “transmission model” of learning (also known as instructivism) by which factual knowledge is transmitted from teachers or instructors to students through lectures, textbooks, and other methods. LS sees the transmission model as effective mainly for declarative knowledge acquisition or the learning of information but limits opportunities to practice applying the knowledge in new contexts, communicating it in complex ways, or using it to solve problems or develop creativity.

After two decades of research and for the above reasons, Hong Kong and Shanghai, two high performing school systems, moved from the transmission model 10 years ago and instituted models derived from the science of learning. Based on how students learn 21st century skills, the models have incorporated the following nine points:

- Make it relevant to the students by leveraging concepts from Project Zero’s *Teaching for Understanding* (http://www.pz.gse.harvard.edu/teaching_for_understanding.php) and Gardner’s theory of multiple intelligences (Gardner, 1991)
- Teach through the disciplines by situating learning activities in “doing” the discipline
- Simultaneously developing lower and higher-order thinking skills
- Encourage learning transfer by applying the knowledge and skills gained in one discipline to another
- Teach students to learn how to learn
- Address misunderstanding directly by understanding students’ models of domain, making them explicit and correcting them
- Promote teamwork as a process and outcome through collaborative and cooperative teaching models
- Make full use of technology to support learning
- Foster students’ creativity (asiasociety.org, 2013).

**Deeper Learning and Learning Outcomes**

As illustrated in the Asian models, LS is also concerned with how to foster deeper learning, or developing durable, transferable knowledge that can be applied to new situations. This is also thought of as the ability to use prior learning to support new learning or problem solving in relevant contexts. Deeper learning is behind the idea that “Success in work and life in the 21st century is associated with cognitive,
intrapersonal, and interpersonal competencies that allow individuals to adapt effectively to changing situations rather than to rely solely on well-worn procedures (Pellegrino & Hilton, 2012).”

Learning outcomes are the end products of formal learning and designed learning experiences found in education and training. Traditionally they have been known as learning objectives and are what the learner should be able to do to after instruction. These are also known as performance objectives describing performances to be exhibited by learners before they are considered competent (Rothwell & Kazanas, 1998). These definitions are found more in the instructional design literature stemming from the instrucivist tradition rather than from LS. However, whether objective or outcome, they are classified by the type of learning desired by a learning taxonomy. As they are hierarchical, outcomes are described as lower-order or higher-order outcomes. They are used as a basis for designing learning activities and assessments and are associated with “doing” verbs. For example, assessing for a knowledge level outcome would be the learner “remembering” something. This is shown in Figure 1 below.

The design of learning environments to foster higher-order learning outcomes (also known as higher-order thinking skills) and assess them is another important aspect of LS. To design instruction and learning environments that are effective, the design must describe what it is that the learner is expected to learn. These are thought of as learning outcomes and also as learning targets (Nitko, 2004). They are typically described in models previously described as learning taxonomies. In these models, learning states (outcomes) are listed and described and the hierarchy is thought to determine an order of precedence in their acquisition. However, this view is not always taken, especially when acquiring learning is described as higher-order outcomes. The design and fielding of SimCalc by Jim Kaput (Balacheff & James, 1997) (http://www.kaputcenter.umassd.edu/products/software) was successful at teaching calculus to students who had not taken any algebra courses, claiming that conceptual calculus outcomes are not situated hierarchically in the typical math curriculum. Another aspect of learning outcome taxonomies is their organization into learning levels and domains. The most familiar may be the three domains described by Bloom—cognitive, affective, and psychomotor (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) of which the cognitive is the most familiar and utilized.
Creativity, innovation, critical thinking, problem solving, decision-making, learning to learn, and metacognition are viewed as 21st century competencies and skills and are typically associated with deeper learning and higher-order outcomes in education and training (Gallagher, in press). Learning taxonomies usually include problem solving, critical thinking, and even metacognition as higher-order outcomes. Cognitive competencies such as cognitive adaptability (considered critical to general adaptability) may be honed through the acquisition of these higher-order outcomes (Gallagher & Prestwich, 2013). Higher-order outcomes are also inherent in the definitions of 21st century skills and competencies and seen as synonymous with deeper learning. Pellegrino and Hilton (2012) define 21st century competencies as both domain-specific content knowledge and the procedural knowledge of how, why, and when to apply it. The Rand report *Teaching and Learning 21st Century Skills: Lessons from the Learning Sciences* identifies critical thinking, problem solving, agility, and adaptability as important components of 21st century skills (Saavedra & Opfer, 2012). 21st century skills are also seen as a global imperative by the U.S. Department of Education to prepare students for the workforce (“21st Century Skills: A Global Imperative | ED.gov Blog,” 2012).

A review of studies in 2004 of the relationships between student performance on large-scale measures and instruction emphasizing higher-order thinking, projects, and multiple-solution problems, reported clear evidence associating this type of instruction with higher scores. On the National Assessment of Educational Progress (NAEP) and the Trends in International Mathematics and Science Study (TIMSS), emphasizing reasoning in mathematics and science was associated with higher scores at all grade levels tested. In reading, teaching for meaning was also associated with higher NAEP performance (Wenglinsky, 2004 in Brookhart, 2010).

**Learning Environment Design**

Learning environments suggest a place and space for learning to take place. Typically this is in a school, classroom or other physical location. However, technology currently allows the concept of a learning environment to extend to the virtual or online “space.” They are thought of as support systems that organize the conditions in which humans learn and accommodate unique needs and relationships for effective learning. Learning environments are considered the structures, tools, and communities that facilitate the acquisition of skills and knowledge (Partnership for 21st Century Skills, 2013). Learning environment design is a prominent topic of research within the LS community. Learning environments are thought of as developed to foster specific types of learning or for developing specific types of learning outcomes. They are based on design principles such as case-based reasoning and how knowledge is integrated (theorizing and testing the nature of knowledge) to environments supporting social interaction using technologies such as computer-supported collaborative learning and mobile learning (Sawer, 2006b). Technology infrastructure for digital or online learning environments are chosen to support tenets of design including those for deeper learning, assessing for understanding and social learning as well a personalized and individualized view of learning such as digital cognitive tutors.

**Assessing for Understanding**

Assessment design in LS is thought of in the context of both teaching and learning and how to improve the processes of both. Assessments targeting deep understanding are well planned using cognitive models guiding the collection of and reasoning about evidences of understanding. Assessments in this mode are
usually embedded in design research models or even action research designs by teachers and may be cyclically employed to understand what the learner is learning and how the teaching process is working (Carver, 2006). LS is not as concerned with standardized or traditional testing for this purpose as it is seen as limited in capturing the data needed to identify true understanding. Standardized testing is even criticized within the LS community because it is seen as engendering instruction that focuses attention on teaching concepts and procedures superficially emphasizing low-level skills, factual knowledge, and the memorization of procedures (Norton & Wiburg, 1998).

Moving beyond research, models exist for assessing higher-order learning outcomes for teachers and instructors. Doing this requires understanding of basic assessment design and the commitment to developing high quality assessments. By identifying the specific learning targets, thinking through the evidence required to show the understanding of the learning target, and developing assessment items that reflect that evidence, formative and summative assessment can be constructed and used by teachers, instructors, or instructional designers (Brookhart, 2010; Nitko, 2004). Although these assessments can take the form of multiple choice or short answer tests, it is typically thought that other alternative assessment types perform better for this purpose.

**Authentic Assessments**
In line with constructivism and deeper learning, learning assessments should ideally occur in a situated environment and authentic context which is described as authentic assessments. What is typically assessed in most e-learning experiences are mostly low-level skills, factual knowledge, and the memorization of procedures typically described as lower-level learning outcomes. The concept of authentic assessment allows assessment to occur in rich context and problems facilitating the assessment of higher-order learning outcomes. Alternative assessment (also known as performance-based assessment) is another name for authentic assessment. In the literature there are nuances between these labels but for most, they are thought of as basically the same. All are variants of performance assessment and all require students to generate rather than choose a response (Norton & Wiburg, 1998). All also use approaches that include the use of open-ended assessment tasks calling upon students to apply their knowledge and skills to create a product or solve a problem and are designed to elicit evidence of attaining higher-order learning outcomes. Authentic assessment specifically refers to an “assessment that both mirrors and measures students’ performance in ‘real-life’ tasks and situations (Hart, 1994 in Norton & Wiburg, 1998) or ‘authentic tasks’ that involve the application of combined knowledge and skills in the context of an actual project (Pelligrino, Chudowsky, & Glaser, 2001).” However, critics of these types of assessments question the alignment between the developers’ goals and intentions and what is actually being measured; and also that the tasks did not actually engage students in the complex thinking tasks desired (Pelligrino et al., 2001).

**Adaptive Expertise**
Another area of interest in LS is the development of adaptable expertise. This is the development of flexible knowledge and attitudes that facilitate effective navigation and problem solving across a variety of settings and tasks. Adaptive expertise is strongly viewed as necessary for learning challenges across cultures (Nasir, Rosebery, Warren, & Lee, 2006). Routine experts can solve familiar problem types quickly and accurately but do not perform as well on novel problems. Although routine and adaptive experts
continue to learn throughout their lifetimes, routine experts develop a core set of competencies that they apply with greater and greater efficiency. Adaptive experts evolve their core competencies, expanding their expertise as needed or as desired. They may even be able to invent new procedures derived from their current level of expertise (J. Bransford et al., 2000; VanLehn & Chi, 2012). Adaptive expertise began to differentiate from routine expertise beginning in 1986. It became more visible and a part of the LS lexicon with its inclusion in the National Research Council's How People Learn (Bransford et al., 2000) and subsequently in The Cambridge Handbook of the Learning Sciences (Bransford et al., 2006). An interesting finding is that adaptive expertise has been implicated in accelerating future learning, facilitating cross-domain transfer of learning (VanLehn & Chi, 2012).

**Personalization and Individualization**

Directions in LS, especially in the educational and instructional technology sphere, have moved to support individualized learning experiences taking into account the individual’s current level of knowledge, past experiences, and preferences among other factors. Inclusive of but not defined as either individualized learning or differentiated instruction (Demski, 2012), this convergence is commonly known as personalization. Supportive strategies are increasingly prominent factors not just in education, national educational technology policy (U.S. Department of Education Office of Educational Technology, 2010), and global education policy (DfES, 2004) but also in the continued pursuit of effective designs in virtual learning environments (Verpoorten, 2009; Verpoorten, Glahn, Chatti, & Specht, 2011).

Personalization in digital learning environments is, most generally, the adaptation of the learning experience (e.g., learning content, methods, tools, pedagogical strategies) according to dynamic and/or static characteristics and actions of the individual learner (e.g., knowledge level, needs, objectives, preferences, styles, and modalities) in order to help the learner to most efficiently and effectively meet stated learning goals and objectives (Verpoorten, 2009). This is done by identifying ways in which the learner’s needs could be identified and learning content adapted to suit those needs (Essalmi, Ayed, Jemni, & Graf, 2010). Striving to meet the needs of an individual can occur from the perspective of a learning system with an external view of those needs to that of the student allowing for the internal needs to drive an individual path to learning. From the former viewpoint, it may mean an intelligent or cognitive tutor continuously assessing a learner’s knowledge and/or affective states providing relevant learning interventions based upon those states. From the latter viewpoint, it could mean empowering the learner to build his or her own responsive learning environment (Fruhmann, Nussbaumer, & Albert, 2009) with many examples existing in the current movement of massively open online courses (MOOCs).

Technologically, supporting the individual needs of the learner in a digital learning environment involves addressing the needs to combine adaptive pedagogical elements (such as personal learning strategies and learning techniques) with available learning tools and information in the most appropriate way to support that learner (Berthold, Lachmann, Nussbaumer, Pachtchenko, Kiefel, & Albert, 2011. It can also involve utilizing automatic personalization and recommendation methods (e.g., profiling methods used in e-commerce, web mining, information filtering and retrieval, user and group modeling and profiling), both self-contained within the learning system and drawing information from outside web behaviors (including web browsing history and user-generated social media content) (Ammari, Lau, & Dimitrova, 2012; Jemni,
Learning Science Report IDA/ADL 08/30/2013

Khribi, & Nasraoui, 2007; Vassiliadis & Stefani, 2012). Technologies supporting these concepts are embedded in the history, current work, and vision of the Advanced Distributed Learning Initiative.

**Advanced Distributed Learning (ADL)**

The Advanced Distributed Learning (ADL) vision and eventually its establishment as an OSD initiative was based on a combination of early empirical findings such as those discussed elsewhere in this report and a projection of trends that were evident in the 1980s. These trends included the continuing effects of Moore’s law, establishment of a global information infrastructure, developing capabilities for natural language understanding and interactions, object oriented architectures and applications, and hand-held or worn computing capabilities. Combined with intelligent tutoring capabilities they pointed to a future of instruction (and performance/decision aiding) provided through individually tailored human-computer dialogues drawing data and information as needed from the global information grid and available on-demand – anytime, anywhere – in real time. This remains the basic vision of ADL – despite its tendencies to concentrate on programming and engineering issues – and the development of what has come to be known as a personal assistant for learning (PAL).

**Accelerated Learning through Individualization**

Through strategies that allow for personalized and individualized learning, research has shown that learning outcomes can be profoundly improved. Effectiveness is also evident in the concept of accelerated learning affecting the time for students to reach competency in stated goals and objectives. The substantial differences found among individual learners are as well established by experimental psychology as they are by casual observation. Some indications of these differences can be found in time to learn. Example findings on the time it takes for different students to reach the same instructional objectives are:

- Ratio of time needed by fastest and slowest students to reach mathematics objectives: 4 to 1 (Suppes, Fletcher, and Zanotti, 1975; 1976)
- Overall ratio of time needed by fastest 10 percent and slowest 10 percent of K-8 students to reach objectives in a variety of subjects: 5 to 1 (Gettinger, 1984)
- Ratio of time needed by fastest and slowest undergraduates in a major research university to learn a programming language: 7 to 1 (Private communication, Corbett, 1998)

That there are differences among students in the speed with which they learn is not surprising, but the magnitude of the differences is. In a classroom with more than 12 learners, there are likely to be some students who are lost and others who are wasting their learning time waiting for others to catch up.

Although the speed with which different students reach instructional objectives is not independent of ability, research has found it most directly keyed to prior knowledge (Dochy, Segers, & Buehl, 1999). Students in military education and training bring with them a wide variety of backgrounds and life experiences – wider than that found among K-12 students. Adjusting the pace of instruction to their individual needs may be especially important.
The challenge this diversity presents to classroom instructors is daunting. Typically they focus on some students and leave the others to fend for themselves. This practice is common in training settings where the primary task is to enable as many learners as possible to reach a specified threshold of knowledge and skill. Technology alleviates this difficulty by allowing each learner to proceed as rapidly or as slowly as needed.

The degree to which individualization matters is to some degree addressed by studies comparing individualized branching with fixed-content, linear sequencing. Two early studies used computer-controlled videodisc instruction to make this comparison (Fowler (1980); Verano (1987). Fowler compared branched presentations with linear instruction in which precisely the same materials were held to a fixed-content, linear sequence. She reported an effect size of 0.72 (roughly, an improvement from the 50th to 76th percentile) in ability to operate and locate faults on a movie projector. Verano compared an interactive, adaptive, branching approach with a strictly linear approach to teach beginning Spanish. He reported an effect size of 2.16 (roughly, an improvement from the 50th to 98th percentile) in end-of-course knowledge. These two studies, among others, suggest that individualization of sequence and content matters, perhaps a great deal (Fletcher, 1991).

One of the most stable findings in comparisons of technology-based instruction with conventional instruction using lecture, text, and experience with equipment concerns instruction time savings. These findings are presented in Table 1.

<table>
<thead>
<tr>
<th>Study (Reference)</th>
<th>Number of Studies Reviewed</th>
<th>Average Time Saved (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orlansky and String (1977)</td>
<td>13</td>
<td>54</td>
</tr>
<tr>
<td>Military Training</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fletcher (1991) (Higher Education)</td>
<td>8</td>
<td>31</td>
</tr>
<tr>
<td>Kulik (1994) (Higher Education)</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>Kulik (1994) (Adult Education)</td>
<td>15</td>
<td>24</td>
</tr>
</tbody>
</table>

As the table shows, Orlansky and String (1977) reported that reductions in time to reach instructional objectives averaged about 54 percent in their review of technology-based military training. Fletcher (1991) reported an average time reduction of 31 percent in 6 assessments of interactive multimedia instruction applied in higher education. Kulik reported time reductions of 34 percent in 17 assessments of technology used in higher education and 24 percent in 15 assessments of adult education (Kulik, 1994). Each of these reviews covered different sets of evaluation studies comparing technology-based instruction to conventional classroom instruction involving lecture, texts, and perhaps laboratory examples. Moreover, time-savings were not their goal. Most of them reported time-savings as an afterthought once they had assessed effectiveness.
Time-savings through the use of this technology do not appear to be achieved at the cost of instructional effectiveness. A review of 233 assessments comparing pre-1992 computer-assisted instruction with classroom instruction found an improvement of 0.39 standard deviations in learning achieved through the use of the technology (Fletcher, 1997).

The 53 studies of early computer-assisted instruction summarized in Table 1 suggest that it could shorten technical training time (i.e., costs) by about 30 percent. These savings would be achieved in a straightforward way through individualization – especially and most simply by minimizing material that learners, with their varying learning rates and prior knowledge, already knew and emphasizing what they had yet to learn. At the time, the DoD was spending about $2,108 million in pay and allowances for individuals in residential specialized skill training. Given the assumption of 30 percent time savings an analysis by Angier and Fletcher (1991) found that if 30 percent of the specialized skill trainees saved 30 percent time in training through the use of computer technology, the savings would amount to $368 million per year or 17% of then current pay and allowances alone. If 40 percent of the trainees saved 30 percent of their time in training, the savings would amount to $491 million per year, or 23 percent of current pay and allowances. These savings only consider trainee pay and allowances. Other training costs are not considered. Moreover, the more rapid graduation of individuals into duty stations would also affect total manpower requirements for specialized skill occupations, which was not considered in this analysis.

In addition to the aforementioned studies, Arizona State University embedded mass personalization technologies in an online developmental math course that continually assesses students’ math proficiency and adapts accordingly, presenting each student with a personal learning path. After two semesters of use with over 2,000 developmental math students at ASU, withdrawal rates dropped by 60% and pass rates went from 64% to 75%. Forty-five percent of the students finished the course four weeks early (Knewton, 2012).

Activation of Prior Knowledge

An important concept to learning and within LS is the idea that effective learning occurs through building on and engaging existing knowledge and experiences (Sawer, 2006a). This not only means uncovering what a learner knows but the misconceptions that exist. Activation of prior knowledge goes a step further allowing the learner to relate to the content and learn using their strengths. This also gets to the core of individualization and personalization in learning systems. As a basis for understanding the individual and evident within the design of personalized digital learning environments is the work of Howard Gardner (1991). Gardner describes intelligence as a spectrum or profile with specific areas of strength. It was initially a set of seven intelligences: linguistic, logical-mathematical, musical, bodily-kinesthetic, spatial, interpersonal, and intrapersonal. Over the years, additional candidates have been discussed, including naturalist, spiritual, existential, and moral intelligences. Educators view this theory as validating students’ everyday experience which provides a conceptual framework for organizing curriculum and pedagogical practices.
Cognitive and Intelligent Tutoring
A foundational area within LS is that of cognitive tutoring also known as intelligent tutoring. As one-to-one tutoring has been intuitively and anecdotally thought to be a more effective way of teaching than classroom instruction, the work of Benjamin Bloom (1984) and Cohen, Kulik, & Kulik (1982) was seen as providing empirical evidence of effectiveness (Bloom, 1984; Graesser & Person, 1994). However, economically, tutoring was not an option that could scale. This began to change with the advancement of computer hardware and software as well as the introduction of the Internet. Technology began to allow the systematic incorporation of LS advances into the classroom, the testing of associated learning principles, and the adaptation of the technology to the needs of students and teachers. This took the form of computer-aided instruction (CAI) and evolved to intelligent CAI also known as intelligent tutoring systems (Fletcher & Rockway, 1986). These systems began to add guided dialog interactions in teaching electronics troubleshooting, adding intelligent questioning to existing CAI systems, and integrating tutoring strategies with medical expert systems for diagnosis. This evolution began to produce cognitive models of knowledge domains, students, and experts, leading to the term cognitive tutors (Koedinger & Corbett, 2006).

Using intelligent tutoring systems or cognitive tutors, digital learning environments are seen to facilitate the attainment of higher-order learning outcomes within specific domains. Using multiple methods and based on LS approach, they are moving students beyond basic facts, simple procedures and concepts to the attainment of adaptive procedures and more abstract concepts (illustrated in Figure 2).

![Figure 2 Overview of Learning Objectives (adopted from Anderson and Krathwohl, 2001)](image)

Anderson’s research has produced a long lineage of research, prominent researchers in the field, and many successful digital cognitive tutors with their roots in ACT-R. Most of this research and development stemmed from Carnegie Mellon University and the NSF funded Pittsburgh Science of Learning Center (now LearnLab). However another line of cognitive tutor research and development came out of the University of Memphis called AutoTutor. Led by Art Graesser, this line of research has produced many descendent systems that hold conversations with humans in natural language. They are based on computational linguistics that analyze and determine cognitive states through students’ dialogue history (autotutor.org,
2012). AutoTutor is a foundational component of the Institute for Intelligent Systems located at the University of Memphis (http://www.memphis.edu/iis/) (University of Memphis, 2013).

**SHERLOCK Project**

An early implementation of cognitive tutoring for the U.S. Air Force was in the SHERLOCK Project. Initially designed as a practice environment for Air Force technicians (Lesgold, Lajoie, Bunzo, & Eggan, 1986), it taught electronics troubleshooting within a complex system involving thousands of parts. It used a simulation of the complex system with an expert system (coach) that gave advice when the learners reached a problem in the troubleshooting process. It also provided reflection tools allowing learners to review their performances and try out ways to improve. It was determined through field tests in performing the hardest real-world troubleshooting tasks, 20 to 25 hours of SHERLOCK training equaled approximately four years of on the job experience (J. Bransford et al., 2000). This had the potential of producing a profound cost savings for training.

**DARPA Digital Tutor Project**

Another profound application of cognitive tutoring in the Navy is in the DARPA Digital Tutor Project. The Digital Tutor project was initiated by DARPA’s Training Superiority program and continued under its Education Dominance program. Both programs were created in response to Defense Science Board recommendations for innovative training research. The Digital Tutor serves two general purposes—meeting a Navy operational need and advancing the technology of computer applications in training. It applies principles from cognitive and instructional theories, but its approach is pragmatic and eclectic rather than theoretic. It is an attempt to make the well-noted advantages of one-on-one tutorial instruction scalable, affordable, and readily accessible through computer technology. Its basic strategy is to observe and capture in systematic and specific detail the practice of individuals who are expert in both subject matter and one-on-one tutoring. Based on an extensive analysis of need and criticality, DARPA selected training for the Navy Information Systems Technician (IT) rating for this effort.

Five assessments of the evolving DARPA Digital Tutor (DT) have been performed (Fletcher, 2011; Fletcher & Morrison, 2012). In addition to standard tests of statistical significance, effect sizes were also calculated. An effect size less than 0.20 is considered negligible, an effect size equal to or greater than 0.75 is considered very large. Effect sizes greater than 0.75 are rarely found in research on education or training.

**Summary of three of the Digital Tutor's assessments**

Assessment Two (IWAR 1) compared the knowledge and skills of human-tutored IT graduates with Sailors who had an average 7.2 years of IT experience in the Fleet. Both assessments showed substantial differences in favor of tutored graduates in IT troubleshooting and IT knowledge, with some effect sizes in excess of 2.00.

Assessment Four compared the knowledge and troubleshooting exercise skills of computer-tutored DT students who had completed the 7 weeks of the Digital Tutor then available with those of graduates who had completed 19 weeks of newly revised classroom IT training. On a knowledge test, DT students outscored graduates of the revised “A” school training with an effect size of 1.95. They also outscored
their instructors with an effect size of 1.35. In troubleshooting, the DT students outscored the “A” school graduates with an effect size of 1.90.

Assessment Five (IWAR 2) concerned the completed 16-week version of the Digital Tutor. It compared the skills and knowledge of DT graduates with those of graduates from the 35-week IT Training Continuum (ITTC) course and those of Fleet ITs with an average of 9.1 years of experience. Among other tests (Review Board interviews, system design and building, security) assessment was conducted with 2½ days of troubleshooting and knowledge testing.

DT graduates outscored the Fleet ITs in troubleshooting with an effect size of 0.85 and ITTC graduates with an effect size of 1.13. The DT teams solved 74 percent of the problems they attempted compared with 51 percent for the Fleet teams and 38 percent for the ITTC teams. DT graduates also solved three of the problems classified as “very hard”, none of which were attempted by Fleet or ITTC participants. In their solution attempts, the probability that DT teams would leave a harmful action in the system was 0.14 compared to 0.41 for Fleet and 0.33 for ITTC participants.

On the Knowledge Test, the DT graduates outscored the Fleet ITs with an effect size of 4.30 and ITTC graduates with an effect size of 3.38.

The greater efficiency, absence of harmful errors, and ability to solve problems at high levels of difficulty demonstrated by DT students suggest both monetary and operational returns of substantial value to the Navy. The Digital Tutor cost $40-50 million to develop. Comparing 16-week DT students graduates with the capabilities of at least 7 years Fleet experience with those who must develop their capabilities over a period of 7 years’ experience and on-job-training suggests annual net present value savings in the Fleet in excess of $350 million.

The design of the Tutor is likely to be a significant advance in the development of training overall and of instructional technology in particular. In addition to a careful front-end analysis to determine training objectives and equally careful selection of expert tutors to ‘clone’, it employs a spiral curriculum approach focused on solving authentic, situated problems.

**Transfer of Learning**

When knowledge is applied in new situations and contexts it exhibits transfer. Transfer of learning is the ideal or ultimate goal of training and education – applying what was learned across situations, domains, and contexts. Learning that is transferable includes not only the content knowledge or the “what’s” but also the procedural understanding of when, why, and how to apply the content knowledge. The process through which transferable knowledge develops is known as deeper learning (Pellegrino & Hilton, 2012). All learning requires transfer from previous experiences including initial learning involving transfer of previous experiences and prior knowledge. Positive transfer requires the activation of the learner’s prior knowledge (J. Bransford et al., 2000). Cognitive scientists discovered in the 1980’s that children can generalize what they learn to a broad range of contexts when they acquire deep knowledge rather than surface knowledge and when they learn how to apply it in real-world social and practical settings (Sawer, 2006a). In the training sphere, when transference of training occurs, it is more effective. For example pilots must be able to solve problems in the air similar to those they have come across in simulators and
instruction manuals (Driscoll, 2000). Transference also has the potential of better efficiencies and cost savings by being able to learn more quickly across contexts and domains with similar situations.

**Learning Science Research Methods**

The cross disciplinary and diverse nature of LS has driven a convergence and emergence of new research methods and approaches. The traditional quantitative, qualitative, and mixed methods in the form of experimental, evaluative, ethnographic, and developmental designs are present but the field has since expanded to include design-based research models over the last decade. Design-based research uses a variety of methods with the goal of understanding the teaching and learning process and theory building through the cyclical act of designing, implementing, and evaluating learning interventions or learning environments (Middleton, Gorard, Taylor, & Bannan-Ritland, 2008).

**Example Learning Science Related Funding and Work**

Over the past year and a half, the National Science Foundation (NSF) has spent $50 million advancing the science of cyberlearning or advancing the science of learning in a technology-rich environment. Having expired by 2013, the program’s intention was to integrate advances in technology with advances in what is known about how people learn in order to:

- better understand how people learn with technology and how technology can be used productively to help people learn, through individual use and/or through collaborations mediated by technology;
- better use technology for collecting, analyzing, sharing, and managing data to shed light on learning, promoting learning, and designing learning environments; and
- design new technologies for these purposes, and advance understanding of how to use those technologies and integrate them into learning environments so that their potential is fulfilled (NSF, 2012).

Another NSF program called Developmental and Learning Sciences (DLS) supports fundamental research to increase understanding of cognitive, linguistic, social, cultural, and biological processes related to children’s and adolescents’ development and learning. Research supported by this program is anticipated to add to the basic knowledge of how people learn and the underlying developmental processes that support learning.

Unlike the Cyberlearning program, DLS is currently active. Among the research topics supported by DLS are: developmental cognitive neuroscience; development of higher-order cognitive processes; transfer of knowledge from one domain or situation to another; use of molecular genetics to study continuities and discontinuities in development; cross-cultural research on development and learning; and the role of cultural influences and demographic characteristics on development (NSF, 2013).

Other federal funding sources include various programs from the U. S. Department of Education Office of Educational Research and Improvement (OERI) and the Institute of Education Sciences (IES) with ongoing funding opportunities for topics related to LS.
Examples of work funded by NSF or the Department of Education

**Institute for Intelligent Systems (IIS) (University of Memphis)**

Led by Andrew Olney, the interim director, the IIS’s stated mission is to advance the state of knowledge and capabilities of intelligent systems, including psychological, biological, and artificial systems. They accomplished this by using an interdisciplinary approach that brings together researchers from many different research areas in the cognitive sciences, including biology, communication sciences & disorders, computer science, education, engineering, linguistics, philosophy, physics, and psychology.

The IIS has received nearly $30 million in external funding over the last decade and has partnered with numerous universities and corporations. Researchers affiliated with the IIS have received funding from the National Science Foundation, the National Institutes of Health, the Institute of Education Sciences (IES), and numerous other agencies and corporate partners (IIS, 2013).

**Learning Sciences Institute (Arizona State University)**

This is a multi-disciplinary community led by Sasha Barab with funded work through IES and NSF in educational gaming, literacy, and educational technology development with stated foci of:

- Learning with adaptive technology and games;
- Learning in dyads or small groups, with students interacting face-to-face, with a conversational agent, or virtually;
- Designing methods for encouraging effective active learning activities in classroom settings;
- Exploring issues related to learning at scale;
- Pursuing new forms of embedded assessment, such as data mining and automated grading systems;
- Understanding the role of motivation, coping skills, and resilience in sustaining interests in learning (http://lsi.asu.edu).

**Sackler Institute (Cornell/Vanderbilt)**

Current research by Bruce McCandliss (Vanderbilt) and others through the Sackler Institute indicate that FMRI (functional magnetic resonance imaging) studies have promise in detecting the success of interventions and how educational learning experiences reshape brain networks supporting basic cognitive skills (http://news.vanderbilt.edu/2011/05/bruce-mccandliss-educational-neuroscience-how-education-shapes-brain-development/).

**LearnLab (Formerly Pittsburgh Science of Learning Center)**

Directed by Ken Koedinger, the LearnLab was initially funded as an NSF science of learning center with Carnegie Mellon University but has since moved out from public funding. Its research thrusts are organized by:

- Cognitive Factors
- Metacognition and Motivation
- Social Communication
- Computation Modeling and Data Mining (http://www.learnlab.org/)
Other than the examples given above, there are a diverse group of researchers within the LS field, including several with lengthy funding portfolios from federal agencies including the NSF, the Department of Education, and the DoD. These researchers include Roger Shank (formerly Northwestern), Roy Pea (SRI/Stanford), Paul Cobb (Vanderbilt), Jim Pellegrino (Northwestern), Richard Lesh (formerly Purdue), Eamonn Kelly (GMU), Brenda Bannan (GMU), Chris Dede (Harvard), Marsha Linn (UC Berkeley), Andrea diSessa (UC Berkeley), Janet Kolodner (Georgia Institute of Technology/NSF Program Officer), Kurt VanLehn (ASU LSI), plus many others not explicitly mentioned.

**Learning Retention**

Learning retention is the ability to maintain the availability of acquired knowledge so that it may be accessed for use at a later time (Driscoll, 2000). The retention of knowledge, skills, and abilities is dependent on many variables. For example, long-term retention of procedural skills beyond initial learning or training is influenced by a host of factors including complexity of the task, use of job or performance aiding, time limits and stress, individual aptitude, and amount of original learning (Wisher, Sabol, & Ellis, 1999). It is important to consider how the knowledge was learned or the type of learning outcome that was achieved. This is related to the type of and end use of the knowledge. For example, what in training are the types of end tasks to be trained? According to Wisher et al. (1999), cognitive, procedural, and perceptual-motor tasks require different types of knowledge and are retained and stored in different areas of the brain. For example, cognitive tasks depending on cognitive abilities such as spatial working memory capacity and executive planning depend heavily on the frontal lobe while recall of verbal knowledge has been shown to occur in the left (usually) neocortex. Sherry & Schacter (1987 in Pinker, 1999) described two primary memory systems: the semantic and the episodic. They suggested that there are two natural memory systems for different purposes. The episodic memory is autobiographical storing memories of specific events as in the first time you learn to ride a bicycle. Semantic memory is called generic-knowledge memory and refers to all of the general information stored in memory that can be recalled independently of how it was learned. For example, you may not be able to recall all of the information surrounding the first time you learned to ride a bicycle but you remember the skill (Driscoll, 2000; Pinker, 1999). These two basic systems were expanded to five (Schacter & Tulving, 1994 in Driscoll, 2000) described as procedural, perceptual representation, semantic, primary, and episodic. Subsequently, a taxonomy of the memory systems very similar to that described by Schacter & Tulving was developed mapping each system to a specific section of the brain and essentially dividing the systems into declarative (explicit) and non-declarative (implicit) memory (Driscoll, 2000).

To illustrate how memory is thought to work, a good example is the cognitive information processing model (Figure 3). Input is received from the senses into sensory memory. Through the processes of attention and pattern recognition, it is stored as information in the short term or working memory. During this process, gathered sensory information may be discarded or quickly forgotten. Short-term or working memory is where information not lost in transmission is acted upon and manipulated or interpreted using retrieval from long-term memory. Working memory however is limited in its storage capacity and duration
of storage\(^1\). For movement into long-term memory, the processes of rehearsal and encoding are used. Long-term memory is the memory described above as being mainly semantic or episodic.

During these processes, loss or decay – forgetting – can occur at virtually any step. Theories of forgetting are important as they describe the impedance of memories forming in long-term memory and can describe why learning is not occurring. Three aspects of forgetting are the failure to encode, the failure to retrieve, and interference. Failure to encode means that the information was never initially learned and cannot be retrieved. Metacognitive strategies and self-regulation of the learning process can help alleviate this issue. Failure to retrieve is the inability to access previously learned information. As retrieval is very much influenced by the context of encoding, this suggests that instruction should use many contexts and situations for retrieval “hooks.” Related is the concept of state-dependent learning which indicates that recall is best in the same situation in which it was learned. This provides an argument for authentic assessment previously discussed. A third aspect is that of interference. As its name implies, this refers to other events or information getting in the way of effective retrieval. This argues for self-paced relevant learning experiences that allow learners to focus on the goals of instruction, which is especially relevant for adults (Driscoll, 2000).

For retention, episodic memory is considered by many to be the most effective for recall and the most durable. Research using word pairs and EEG has shown retrieval is better in episodic memory compared to semantic (deep encoding) or implicit memory shallow encoding (e.g., unrelated facts) (Wieser & Wieser, 2003). For learning, knowledge situated in personal experience that is meaningful to the learner is more apt to be retained providing the necessary “hooks” for retrieval. In contrast, semantic memory recall is considered not as effective for long term retention. From the viewpoint of LS, it will quickly decay. This may be due to lack of context and state dependency resulting in retrieval failures.

---

\(^1\) In a classic study in 1956, it was demonstrated that 7 ± 2 numbers could be recalled in a digit-span test (Miller, 1956).
Another way to increase retention is through repeated practice, or overlearning. Often occurring in instruction or environments with frequent testing, continued practice and retrieval of information can produce long-term retention. In 2006, subjects in a repeated test condition outperformed those in a repeated study condition and a control condition (Karpicke & Roediger, 2007). This is not a new phenomenon as Wisher et al. (1999) pointed out; tasks that are overlearned turn out to be highly resistant to decay.

Not taking into account how the knowledge was acquired, Wisher et al. (1999) found that different types of learned tasks have different rates of retention decay. For example, research on memory for cognitive tasks showed a moderate rate of decay with a relatively small amount of forgetting occurring for up to a year after learning. This is not a surprise from a LS perspective as interpretation, decision making, and troubleshooting are all extremely contextual and situated within a domain. Wisher et al. (1999), also discuss remarkable retention rates for what they call “school knowledge” or knowledge learned in school that may last for a lifetime. The retention factor in this case is dependent of whether or not a memory test requires the knowledge to be recalled or recognized with recognition decay performing much better than pure recall. The lasting effect of school knowledge is also not surprising as much of school experience may be embedded in episodic memory tied to other events both good and bad that could facilitate this recall. Some procedural tasks, however, do not seem to fare as well. Procedures require the recall of both steps to be done and possibly their order. Depending on their type (continuous execution or discrete procedure) they may either have very good retention or very bad. Continuous execution tasks such as typing, aircraft flight control, target tracking, or even bike-riding, show virtually no skill loss for periods as long as two years without practice. However, procedures requiring discrete steps such as clearing an M16 rifle have shown to have low retention with pronounced declines after only five or six weeks (Wisher et al. 1999). Unfortunately, the type of instruction or learning that occurred when acquiring these skills in the studies is not identified. LS would argue that the way in which they are learned matters to retention with highly contextual and situated learning resulting in a higher probability of retention. From a neuroscience perspective, this would mean that instruction should ensure that the episodic memory is activated providing a much richer means of retrieval.

Interventions that Support Better Retention

Retention of learning begins with how the knowledge or skill was learned. Learning interventions are different methods, strategies, or modes of learning design, instruction, or learner support designed to instill learning goals or produce specific learning outcomes. Based on the current literature covering studies involving learning retention, several effective learning interventions have been identified. They center on interventions designed for active learning and deep conceptual learning facilitating the encoding into episodic memory. However, they also include interventions embedding performance aiding or performance support concepts. This is not surprising as one of the best times for learning to take place is at the moment of cognitive or performance breakdown within a task.

Problem-based, authentic, situated, and active learning

- Multiple studies have shown that the use of problem-based and generative learning principles (e.g., problem-based learning, situated learning, active learning, tool usage to significantly increase retention of knowledge and skills in a variety of different contexts. For instance,
Subramanian, Timberlake, Mittakanti, Lara, and Brandt (2012) found that the use of situated cognition, authentic practices, and case-based learning embedded in a web-based learning environment resulted in a significant improvement in medical student learning and retention as compared to the traditional didactic lecture format (long-term improvement from baseline knowledge was 10.9 ±4% for the control group compared to 26 ±3% for the experimental group).

• Additionally, Anderson (2007) demonstrated in the context of teaching leadership theory in urban high school agricultural programs, problem-based learning resulted in higher retention than teacher-guided learning after a period of three months \((F = 7.75, p < .05)\).

• Finally, active learning strategies involving social negotiation, simulations, collaboration, and problem-based learning was found to engender better conceptual learning than traditional instruction when teaching Cantor Set Theory (Narli, 2011) thought to lead to better retention \((\chi^2=15.61, sd=2, p<0.05)\).

Teaching for conceptual understanding and deep conceptual knowledge

• A conceptual understanding of the underlying foundations of the subjects being taught has also been found to increase knowledge retention. For instance, students in an inquiry-based class were shown to have retained a greater amount of conceptual mathematical knowledge dealing with differential equations after one year than students in traditionally taught classes \((F=7.24, p=.01)\) (Kwon, Rasmussen, & Allen, 2005).

• Also, University of British Columbia students taught with non-traditional methods emphasizing conceptual knowledge retained more conceptual quantum mechanics when compared to average student retention (Deslauriers & Wieman, 2011).

Repeated retrieval of information

• Continuing practice or overlearning, i.e., the repeated retrieval of information, has been shown to produce long-term retention. Karpicke and Roediger’s (2006) study compared the retention of students who were repeatedly tested during the learning process (repeated test condition), those who repeatedly studied the previously learned information (repeated study condition), and those who did neither (control condition). Students in the repeated test condition had greater retention than those in the repeated study or standard conditions, indicating that being forced to repeatedly retrieve previously learned information leads to greater retention. Studies have also shown that high initial aptitude of trainees can be predictive of a tendency toward overlearning (Wisher et al., 1999), so at an organizational level, selecting high-aptitude trainees can result in higher degrees of overall retention in the trainee population.

Performance aids

• Bink and Cage’s (2012) study showed two different performance aids, specifically geared toward the needs of high and low performers, respectively, to each improve the retention of map reading skills in Army Initial Military Training (IMT) after three weeks \((F = 24.45, p < .01)\).

• In the context of a cost-benefit analysis, an electronic performance support system (EPSS) for the customer service application used by one of the world’s seven largest utility companies (THEO-EPSS) was found to provide useful knowledge to novice users and allow them to perform well on tasks that would normally require initial training, showing that THEO-EPSS can reduce initial
training (Desmarais, Leclair, Fiset, & Talbi, 1997). Novices and experienced users completed 15 tasks on finding information with the EPSS present, without the EPSS, and repeatedly for practice effects. Subjects performed evenly on the first condition, novices failed in the second condition, time was reduced from 30 to 15 minutes per task in the third condition.

- Users of embedded or intrinsic performance support systems (incorporated directly into the work interface) attain organizational competency levels at a very early stage in preliminary training scenarios (Gal & Nachmias, 2011). Additionally, the use of embedded (intrinsic) performance support with or without training was shown to improve performance and reduce the time to complete a task compared to training only: Performance – $F(2,46) = 22.37$, $p<.01$; $\eta^2 = .31$ and Time to perform task – $F(2,45) = 11.2$, $p<.01$; $\eta^2 = .32$. (Nguyen & Klein, 2008).

- Organizational implications for embedding performance support practices may have a large cost benefit. A shift within a large commercial organization from formal learning as a primary training intervention to performance support resulted in an estimated productivity savings of $2.6$ million over the course of 2011 (Lanese & Nguyen, 2012).

### Impact of Learning Science on Sustainment of Skills and Knowledge
Sustainment training is considered training required to maintain the minimum acceptable level of skills proficiency and knowledge levels required for competency. It is used to refresh memory of knowledge and the skills needed to perform tasks. Requirements for sustainment training depend on the type of knowledge and skills, the strategies used to design and conduct initial training to acquire the knowledge and skills, and the rapidity with which the skills decay, and the rapidity with which knowledge and skills can be returned or raised to required levels of performance. These levels are also referred to as a “go” state and the type of training associated to produce this state can also be called refresher training (Wisher et al., 1999). In any case, maintaining this state is dependent on the amount of retention from initial training as well as transfer of training and larger factors such as selection and classification (Zeidner, 1987). Training within the services occurs on an individual or collective basis and is divided between training occurring as residence training or in operational units (Table 2). Who is being trained, how it occurs, and where it occurs (environment) are important factors when considering effectiveness of interventions. Considerations also include the costs of training, when and how refresher training occurs, other learning interventions and the trade-offs in real and opportunity costs associated. This section discusses the how LS affects sustainment training particularly in training transfer, training evaluation, and potential cost models.

### Table 2 Who is Trained and Where it Occurs

<table>
<thead>
<tr>
<th>Where Training Takes Place</th>
<th>Residence</th>
<th>Operational Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who is Trained</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Training Transfer Considerations**

The reason for training is enhanced productivity and operational effectiveness. However carefully a learning environment is prepared, performance in the field will require more than it can provide. Transfer of what’s been learned to what a specific job or task requires will be needed. Realism, authenticity, or fidelity of training remains a perennial topic of discussion.

This discussion generally responds to the intuitive appeal of the argument that transfer is enabled by elements presented in training that are identical to those found in the task or job. In addition to articulating the argument in the first place, Thorndike and Woodworth (1901) additionally asserted that such transfer is always specific and may be keyed to either substance or procedure. Later studies of transfer reinforce their point of view. It seems reasonable to expect task elements mastered in training to be performed with some appreciable degree of success on the job.

The focus on identical elements often leads to an insistence on maximum fidelity in education and training. As a consequence, there is a tendency to provide as many identical elements in training as possible. This would be a viable approach if fidelity came free, but it does not.

As fidelity rises so do costs. High costs can be borne but will also reduce the number, availability, and accessibility of other features and resources that can be routinely provided for training. A solution is to adopt a degree of fidelity that matches specific training objectives. The optimal choice of fidelity keys, therefore, on careful explication of training objectives. The process of basing fidelity on training objectives is usually called “selective fidelity.” The dilemma in this case shifts to the difficulty of specifically identifying relevant training objectives in advance of training.

The identical elements theory of transfer is itself limited. It works well for transfer of procedural tasks that are completed in roughly the same way each time they are done. However, more complex tasks involving problem solving and judgment are not well served by identical element, high-fidelity training. Far transfer, in which the training and transfer settings differ, usually rely on analog transfer, which is not based on surface features, but on the deep structure of problems – those are best trained using capabilities such as those found in intelligent training systems focusing on deep and/or conceptual learning. Experts in many fields solve problems through transfer-enabling analogies they form from an extensive database of solution principles assembled over years of experience. Intelligent tutoring systems
often accelerate the development of expertise by tailoring problems for individual learners in the same way that experts in many fields engage in deliberate practice of competencies they have yet to master (Ericsson, Nandogopal & Roring (2009)).

**Evaluation of Training**

The evaluation of training interventions, which could be extended to performance aiding as well, is an important tool to understanding what is effective in what context leading to decisions affecting not only the effectiveness but the costs and cost trade-offs of the intervention on an organizational level. Many models exist to frame the effectiveness of training, but the most visible and most used model is from Donald Kirkpatrick beginning in the 1950’s but taking off with the publication of his book *Evaluating Training Programs: the Four Levels*, which is now in its 3rd edition (Kirkpatrick & Kirkpatrick, 2006).

As stated, the purpose and goal of military training is enhanced success in performing military operations. This point of view for all training was emphasized by Kirkpatrick (1987) who summarized long-standing ideas about training effectiveness with four levels of assessment shown in Table 3. When applied, these levels of assessment are seen as increasing more difficult to execute and potentially needing an increasing level of resources. These levels might be described as:

Level 1: Surveys. Level 1 assesses the opinions and beliefs of learners, instructors, and administrators. It is not without value, and it is the most readily available and inexpensive form of assessment. However, it tells us more about opinions than about measured training effectiveness.

Level 2: Training Outcome Measures. Level 2 assessments determine if the training achieved its objectives – if it produced the knowledge, skills, and abilities needed to perform targeted tasks to specified standards under specified conditions.

Level 3: Transfer to Job Performance. Level 3 assessments determine if the knowledge, skills, and abilities acquired through training transfer to duty station performance.

Level 4: Benefits to the Sponsoring Organization. Level 4 assessments determine if the acquired knowledge, skills, and abilities enhance organizational effectiveness, productivity, and likelihood of success.

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Evaluation Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Surveys</td>
<td>What were the opinions of the training?</td>
</tr>
<tr>
<td>2</td>
<td>Training Outcome Measures</td>
<td>Did the training achieve its objectives?</td>
</tr>
<tr>
<td>3</td>
<td>Transfer</td>
<td>Did job/work performance improve?</td>
</tr>
<tr>
<td>4</td>
<td>Benefits</td>
<td>Did organizational performance improve?</td>
</tr>
</tbody>
</table>
In the model, Levels 1 and 2 focus on doing things right and Levels 3 and 4 focus on doing the right things.

These four steps are as relevant to military training as they are to civilian applications, and they are applied with about the same frequency in military training as they are in civilian training. Many assessments of military training stop at Level 1, with surveys of the trainees and instructors. Some proceed to Level 2, with end-of-course measures of knowledge and skills provided by the training. A few assessments of military training extend to Level 3, with measurement of on-the-job performance improvements. Even these are frequently limited to opinion surveys of peers and supervisors and rarely involve genuine measurement of job performance.

At Level 4, civilian trainers consider issues such as productivity. Military trainers must consider operational effectiveness, which keys on combat effectiveness. A few Level 4 evaluations of military training have been undertaken to link training inputs with operational effectiveness. For instance, Cavaluzzo (1984) found that a 1-percent increase in flying hours was associated with a 0.5-percent decrease in average bombing miss distance; Hammond and Horowitz (1990) found that a decrease of 10 percent in flying hours increased the probability of defeat in air-to-air combat by 9.2 percent; and Holz, O’Mara, and Keesling (1994) found that the number of tank miles driven was correlated with offensive mission performance \( r = 0.68 \) and defensive mission performance \( r = 0.80 \).

However, Kirkpatrick’s levels need to take a step further (which Kirkpatrick eventually added). Assessment is intended to inform decisions. The hallmark of decision-making is not simply a matter of selecting the most effective alternative, but also accounting for what must be given up to get it – its cost. Unfortunately, cost is rarely considered in education and training assessment, thereby weakening the ability of these assessments to influence decisions. Cost analysis could and should be considered more frequently and more substantively in assessments of military training.

Assessment and Cost Analysis in Training

Training effectiveness research needs to be integrated with cost analysis. Doing so would help answer Admiral Ike Kidd’s famous question, “What is a pound of training worth?” Senior decision makers in Defense must make macro trade-off decisions between material (airplanes, tanks, boats and ships), equipment (spares, fuel, ammunition), and training. The military value of materiel and equipment may be sufficiently obvious, but allocating scarce funds between them and training, lacks significant rigor without explicit measures of value applied to training. Defense executives are often willing to take the value of training on faith, but faith only goes so far when hard budget allocations are at stake. A serious effort to assess the value of training in monetary terms might meet with success and substantially inform decisions about its funding in relation to other investments in Defense. At present, however, both research and analyses to answer the Admiral’s question are insufficient – i.e., effectively absent (Fletcher & Chatham, 2010).

The goal of training research does not end with better training. As a means to an end, training effectiveness needs to address military productivity and effectiveness through the human performance it enhances. Training is not the only contributor to human performance. For instance, personnel selection for military service, classification of personnel into appropriate occupations and careers, ergonomic
design of equipment, and provision of job performance aids all contribute to human performance and the operational effectiveness dependent on it. These four factors directly interact with training and may increase or decrease demands (i.e., costs) of training. The investment trade-offs, however, are poorly understood and coordinated. Systematic investment allocations across all factors intended to produce and sustain human effectiveness are needed. They can and should be enhanced through the use of cost and effectiveness trade-offs intended to help rationalize, if not optimize, allocations of perennially scarce manpower, personnel, and training resources.

Developing a model – a multi-dimensional convex hull -- to determine optimal trade-offs among all these factors (to say nothing of others that might be included) seems impractical at this stage of research and data. However, pairwise trade-off models are possible. They would suggest satisficing balances for investments in these areas, and seem worth pursuing, given the need for them. The ultimate objective of these models would be to ensure that maximum levels of human performance and competencies are made available for military operations through integrated allocations of acquisition, manpower, personnel, and training resources.

Interventions to Improve Sustainment of Skills and Knowledge
Improving retention of skills and knowledge is at the core of sustainment. There are many theories and studies discussing what should occur to maximize two important considerations: retention of learning and transference of learning. Wisher et al. (1999) suggests six methods that could support sustainment and are in agreement with the previous research and evidence presented in this discussion.

- Analyze tasks for their vulnerability to decay and optimize the schedule of refresher training accordingly.
- Increase the number or length or training sessions to maximize “overlearning” of the skills.
- Test often and repeatedly to maximize retention.
- Provide opportunities for spaced practice and repeated practice sessions.
- Use task-oriented training for context and conceptual learning.
- Use peer tutoring to enhance original learning.

The first four methods do not address how learning occurs and may be more of a brute force approach that may run up against increased costs and scheduling issues in their implementation. However, the last two methods (task-oriented training and peer tutoring) provide more of an approach for producing deeper learning or conceptual knowledge and may not be as resource intensive.

Teaching for Deep Conceptual Knowledge
Common to most evidence and theories within LS for maximum retention and transfer is the teaching of deep conceptual knowledge and, where possible, requiring episodic memory encoding. As with education, this practice speaks to the need for ensuring that training is situated, relevant, authentic and social. The example of the Digital Tutor shows how digital systems with social interaction can help this to occur; however, LS offers many methods and models for deep learning. Teaching for deep conceptual knowledge also tends to produce higher-order learning outcomes consistent with 21st century skills and competencies, which may help in the eventual transition to the civilian workforce.
Performance Support or Aiding

The concept of embedding performance support systems within tasks and software systems has been shown to profoundly reduce and, in some cases, eliminate the need for training on these tasks. Not only do these systems enable better performance, but learning occurs allowing the learner to become proficient independent of the system. As mobile technology has become ubiquitous, the use for mobile performance support has increased and has been identified as a prominent need of the community at large through a recent needs analysis performed by ADL for its Mobile Training Implementation Framework (MoTIF) Project. As an example of how far this approach has come, consider Tracking Point’s Precision Guided Firearms (PGFs) utilizing computerized scopes and an iPhone for precision rifle targeting (Hutchinson, 2013). Use of the Integrated Maintenance Information System in the Air Force for maintaining only 3 of many F-16 avionics subcomponents was shown to save the Air Force at least $3 million per year (Teitelbaum & Orlansky, 1996). Using performance support as a primary learning intervention has been shown to save one large organization $2.6 million in costs over the course of one year.

Use of Cognitive Tutors

The use of cognitive tutors is an evolving and often discussed method of teaching deep conceptual knowledge situated within authentic contexts and problems. Currently led by the Digital Tutor but evident in many projects and technologies currently being supported by ADL research projects under the PAL line of research and development, these tutors and tutor technologies may provide tractable ways to instill deeper learning, thereby positively affecting retention and sustainment of skills and knowledge. Problems still to be solved may lie in overreliance of external learner conditions and behaviors for adapting pedagogies and instructional strategies to students’ needs. As learners mature and gather expertise in a domain and in their own understanding of the learning process, systems may need to re-evaluate these conditions allowing more and more learner control in this process. This essentially is the migration from an extrinsic to intrinsic viewpoint of the learner. Examples of where this migration works well are in the concept of personal learning environments exemplified in what we know as MOOCs. Other problems lie in the domain dependence of most tutors and how a successful digital tutor can migrate towards new learning domains which may require different forms of pedagogical content knowledge and possibly different models of the learner which may inhibit scaling. Progress has been made on this front with the introduction of the Generalized Intelligent Framework for Tutoring (GIFT) developed by the Army Research Laboratory (Sottilare, Brawner, Goldberg, and Holden, 2012).

Big Data and Learning Analytics

Recently the term “big data” was developed to refer to the storage of large quantities of longitudinal and transactional data generated through every interaction occurring online. In online learning, including most digital learning systems, records are maintained detailing not only student outcomes but virtually every move made in the online world. These interactions include entries on assessments, discussion board entries, blog entries, wiki activity, and social network activity. These could amount to thousands of events per course or learning experience per student (Picciano, 2012). In addition, log activity is generated in every web page, online game, online simulation, or virtual world interacted with whether specifically in a
managed learning environment or not. Collectively these data amount to an almost unlimited amount of
data collected and stored by various online systems.

Commonly known as data mining, analytics is generically used to define data-driven decision making using big data and learning analytics defines data-driven decision making specifically for learning purposes (Picciano, 2012). As learning occurs through various means in the digital world, it is important to facilitate the collection of learning interactional data from each system generating it ideally in an interoperable way. These sources could be from learning management or course management systems, digital tutor systems, performance support systems, serious games, or full task trainers (e.g., aircraft or tank simulators). Once captured, these data can be stored and, along with other data sources such as competency models or even anecdotally captured data, fed into a learning analytic engine for deciding on various courses of action. These data sources and processes are outlined in the Learning Analytics Flow Model below (Figure 4).

Many areas of concern can be addressed using learning analytics but seven popular categories have been identified which provide an initial lens to view learning using big data. These are:

- Monitoring individual student performance
- Disaggregating student performance by selected characteristics
- Identifying outliers for early intervention
- Preventing attrition from a course or program
- Identifying and developing effective instructional techniques
- Analyzing standard assessment techniques and instruments

![Figure 4 Learning Analytics Flow Model (adopted from Picciano, 2012)]
To accomplish even the previous seven applications, a common method for data capture is needed to ensure that learning interactions are captured in the same way across all organizations and activities. This commonality is referred to as interoperability and is the reason data standards and specifications exist. In the lexicon of instructional technology, this data capture for learning purposes is called learning tracking. ADL has provided for methods to track learning data using the Sharable Object Reference Model (SCORM) which supported the limited tracking of early web-based training applications. However, it was determined through interaction with the training community at large that tracking of robust “big data” for learning was important and ADL has since been developing open source technology called the Experience Application Programming Interface (xAPI) to facilitate this very need. Using xAPI, it is possible to track learning interactions interoperably from virtually any digital data source with online access. In combination with another open source ADL technology called the Learning Record Store (LRS), it fulfills the center two blocks in Figure 4 in a web-services based approach that could be deployed in almost any web-based environment.

Leveraging the big data generated through myriad of learner interaction data in combination with other systems of record managing personnel, force strength, competencies, or other data, may be a cost saving method for understanding training effectiveness, retention issues, refresher cycles or other concerns. In this way, it may be more tractable to get to Kirkpatrick’s Level 3 and Level 4 evaluations potentially moving into the realm of return on investment.

**Conclusion**

This report has discussed learning science and its current topics. It includes examples of learning science research efforts; its basic understanding of learning and retention of skills and knowledge; how these conceptual foundations lead to the design of practical learning environments that yield longer term retention, transfer, and general sustainment of the skills and knowledge they produce; representative examples of these learning environments; ongoing projects to continue research and development in learning science; and assessment issues that arise from learning science, including a practical need for cost analysis and balanced investment trade-offs. Although the topics in learning science fundamentally concern human cognition and memory, applications of technology frequently emerge as affordable and accessible means to implement its findings in practical environments. These applications promise to accelerate learning and enhance human competencies and performance through education training for learners of all ages in all sectors of the economy.

**References**


