

A Cognitive Approach to Instructional Design for Multimedia Learning

Stephen D. Sorden
Northern Arizona University
Flagstaff, AZ, USA

steve.sorden@nau.edu

Abstract

Aimed at both newcomers to online learning as well as experienced multimedia developers, this paper addresses the issue of how to avoid unproductive multimedia instructional practices and employ more effective cognitive strategies. Baddeley's model of working memory and Paivio's dual coding theory suggest that humans process information through dual channels, one auditory and the other visual. This, combined with Sweller's Theory of Cognitive Load and Anderson's ACT-R cognitive architecture, provides a convincing argument for how humans learn, which leads to the question of how multimedia instruction can be designed to maximize learning. Cognitive theory and frameworks like Mayer's Cognitive Theory of Multimedia Learning provide empirical guidelines that may help us to design multimedia instruction more effectively. Mayer argues that the best way to present multimedia instruction is through visual graphics and informal voice narration, which takes advantage of both verbal and visual working memories without overloading one or the other.

Keywords: working memory, multimedia, cognitive load, act-r, production system, dual coding.

Introduction

Cognitive theory is borne from the relatively new interdisciplinary field of cognitive science. Cognitive science studies the nature of the mind by drawing from research in a number of areas including psychology, neuroscience, artificial intelligence, computer science, linguistics, philosophy, and biology. The term **cognitive** refers to perceiving and knowing, and cognitive scientists seek to understand mental processes such as perceiving, thinking, remembering, understanding language, and learning (Stillings, Weisler, Chase, Feinstein, Garfield, & Rissland, 1995). As such, cognitive science can provide powerful insight into human nature, and, more importantly, the potential of humans to develop increasingly powerful information technologies.

This paper addresses the problem that much of what we are currently seeing in multimedia in-

Material published as part of this journal, either on-line or in print, is copyrighted by the publisher of the Informing Science Journal. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission and payment of a fee. Contact Editor@inform.nu to request redistribution permission.

struction may actually hinder the learning that it claims to promote and then discusses possible ways to improve it. I introduce several well-known assumptions of cognitive science, which provide a framework for applying empirical theories of cognition and learning that improve multimedia instruction and assist humans in learning more effectively. The cognitive theories discussed in the paper

include the Theory of Working Memory, Dual Encoding Theory, Cognitive Load Theory, ACT-R Production System Theory, and the Cognitive Theory of Multimedia Learning. Since most instructors have either already been tasked with creating multimedia instruction, or soon will be, this paper is aimed as much at the general practitioner of multimedia instruction as it is the experienced e-learning developer.

Popular forms of multimedia instruction, such as online learning and the more inclusive computer-based training (CBT), have created many new possibilities for education. They provide new ways of delivering content, and they often promote learner-centered environments that can motivate students and add variety to learning. In this environment, instructional units are often accompanied by a liberal use of multimedia that is intended to add excitement to the lesson and hold the learner's attention. However, visual and auditory components that are intended to stimulate rather than educate do not always make for sound instructional design in multimedia delivery and can quickly become counter-productive to learning.

The human mind is limited in the amount of information that it can process (Miller, 1956). Because computer-based training can quickly overwhelm these limited capacities (Sweller, 1988, 1994), it becomes important for the instructional designer to understand the principles of cognitive science and how they apply to effective instructional design for online learning. Concepts, such as working memory, cognitive load, production system theories of knowledge and learning, self-explaining behaviors, and transfer, all become important considerations for the instructional designer who must learn to use technology effectively and intelligently, rather than simply because it is available and seems flashy or exciting.

This is especially relevant as education begins to turn to gaming as the latest innovative technology that some educators claim will revolutionize learning. Proponents of gaming in education, however, should remember that similar predictions were made for mimeograph machines, overhead projectors, movies, radios, television, and the computer, only to produce disappointing results after considerable expenditures of money (Cuban, 1996, 2001). One concern should be that using video games as an educational medium may actually decrease learning in comparison to simply presenting the information in a straightforward manner using text and pictures.

Until recently, much of what we have seen in multimedia instructional design appears to be based more on intuition than empirically-based research. For example, it might seem that an online activity that uses flashy multimedia and game-like strategies to hold a learner's attention is good. The learner is, after all, engaged and his or her attention is fully focused on the activity at hand. Because it is possible and it seems to emulate a tutoring session, why not throw in a talking figure that appears on screen and guides the student through the learning process with jokes and lively gestures? If there is some educational purpose tied to all of the activity on the screen, then, at the very least, some implicit learning must be happening, which, one might argue, is better than no learning at all. But cognitive scientific research and instructional science literature is starting to call some of these assumptions into question (Clark & Mayer, 2002). It is very probable that much of what is occurring under the label of CBT and e-learning is wasted time or less-than-optimal instruction. Research suggests that there is a place for CBT and online learning, but it also cautions us to structure it in a way that efficiently maximizes learning. What is most important is not whether the instruction takes place in a classroom or on a computer screen, but whether empirically-tested strategies for multimedia instruction are employed that facilitate knowledge construction by the learner.

We will look at some of Richard Mayer's recommended guidelines for more effective multimedia instruction, but first let's consider some of the key assumptions that form the basis of cognitive theory in relation to human memory and how we learn, beginning with working memory and its limitations.

Working Memory

Working memory is a concept that grew out of the older model of short term memory (Atkinson & Shiffrin, 1968), which was seen more as a structure for temporarily storing information before it passed to long-term memory. By the late 60's and early 70's, researchers began to question some of the assumptions of short-term memory, however, and a few started to look for more satisfactory explanations. Baddeley and Hitch (1974) eventually proposed a more robust model of short-term memory, which they called working memory. Their model for working memory was a system with subcomponents that not only held temporary information, but processed it so that several pieces of verbal or visual information could be stored and integrated.

Under this model, Baddeley (1986, 1999) proposed that there was a component in working memory that controlled subcomponents or slave systems. This core system, dubbed the **central executive**, was responsible for controlling the overall system and engaging in problem solving tasks and focusing attention. Baddeley theorized that the central executive could transfer storage tasks to two slave systems in working memory, so that the central executive would continue to have capacity for performing more demanding information processing tasks.

These two slave systems eventually became known as the **visuo-spatial sketch pad** and the **phonological loop**. The visuo-spatial sketch pad is assumed to maintain and manipulate visual images. The phonological loop stores and rehearses verbal information, and it has been suggested that it also has an important evolutionary function in that it facilitates the acquisition of language by maintaining a new word in working memory until it can be learned (Baddeley, Gathercole, & Papagno, 1998). More recently, Baddeley (2002) has proposed that it may be necessary to add a third subsystem to his model, known as an **episodic buffer**, which has acquired some of the tasks that were originally attributed to the central executive (now seen as a purely attentional system), specifically functioning as a storage structure which acts as a limited capacity interface to integrate multiple sources of information from other slave systems.

If we accept the concept of working memory in instructional design, the next question we should probably ask is, are there limits to how much information can be processed by working memory, and if so, how can we manage this bottleneck? Cognitive Load Theory states that there is indeed a limit to the amount of information that can be processed at one time, which creates important considerations for multimedia instructional design.

Cognitive Load Theory

Cognitive Load Theory (Chandler & Sweller, 1991; Sweller, 1988, 1994), or CLT, states that working memory is limited in its capacity to selectively attend to and process incoming sensory data. CLT is concerned with the way in which a learner's cognitive resources are focused and used during learning and problem solving, suggesting that for instruction to be effective, care must be taken to design instruction in a way as to not overload the mind's capacity for processing information. The implication for multimedia instruction is that if we only have a very limited amount of information processing capacity in working memory at any single moment, then instructional designers should not be seduced into filling up this limited capacity with unimportant but flashy "bells and whistles" in a multimedia instructional unit.

An example of what this means for multimedia instructional design is that the layout should be visually appealing and intuitive, but that activities should remain focused on the concepts to be learned, rather than trying too much to entertain. This is especially true if the entertainment is time consuming to construct and is complicated for the learner to master. Working memory can be overloaded by the entertainment or activity before the learner ever gets to the concept or skill to be learned.

According to Sweller, content knowledge is organized into **schemas** found in long-term memory, which can be loosely equated to Miller's (1956) concept of a **chunk**, and these schemas control how new information is handled as it enters working memory. Schemas organize simpler elements and can then act as elements in higher order schemas. In other words, as learning occurs, increasingly sophisticated schemas are developed and learned procedures are transferred from controlled to automatic processing. **Automation** frees capacity in working memory for other functions.

In multimedia instructional design, this suggests that a task analysis should be done to break down the skills and information that are needed to learn or perform the educational objective. The multimedia lesson should try to ensure that the learner has sufficiently automated key core knowledge or tasks, before trying to tackle an overall task that may be beyond the learner's current ability range, causing unnecessary frustration and possibly even that the learner drops out of the activity. Readers may recognize features of Vygotsky's *Zone of Proximal Development* and Piaget's concept of *scaffolding* here.

This process of developing increasingly complicated schemas that build on each other is also similar to the explanation given by Chi, Glaser, and Rees (1982) for the transition from novice to expert in a domain as illustrated by De Groot's (1966) study of chess grand masters compared to less able players. One explanation is that the grand masters achieve expert status not necessarily because they processed information any faster than the novices, but because they memorize entire patterns or configurations of chess pieces on the board and employ appropriate strategies based on an overall pattern, rather than the positions of individual pieces. Similar to these memorized chess patterns, schemas increase the amount of information that can be held in working memory by chunking individual elements of information into more complicated elements.

CLT suggests that instructional techniques that require students to engage in activities that aren't directed at schema acquisition and automation can quickly exceed the limited capacity of working memory and hinder learning objectives. In simple terms, this means that you shouldn't create unnecessary activities in connection with a lesson that require excessive attention or concentration that may overload working memory and prevent one from acquiring the essential information that is to be learned. This is an important rule in any form of instruction, but it is an essential rule in multimedia instruction because of the ease with which distractions can be incorporated.

According to Sweller, Van Merriënboer, and Paas (1998), there are three types of cognitive load: intrinsic, extraneous, and germane. The first, **intrinsic cognitive load**, occurs during the interaction between the nature of the material being learned and the expertise of the learner. The second type, **extraneous cognitive load**, is caused by factors that aren't central to the material to be learned, such as presentation methods or activities that split attention between multiple sources of information, and these should be minimized as much as possible. The third type of cognitive load, **germane cognitive load**, enhances learning and results in task resources being devoted to schema acquisition and automation. Intrinsic cognitive load cannot be manipulated, but extraneous and germane cognitive load can.

CLT states that an instructional presentation that minimizes extraneous cognitive load can facilitate the degree to which learning occurs. Chandler and Sweller (1991) demonstrated that one method for reducing extraneous cognitive load is to eliminate redundant text. Mousavi, Low, and Sweller (1995) and Sweller et al. (1998) argued that cognitive load is reduced by the use of dual-mode (visual-auditory) instructional techniques and that the limited capacity of working memory is increased if information is processed using both the visual and auditory channels, based on Baddeley's model of working memory. Intrinsic, extraneous, and germane cognitive loads form an equation in which the sum total of the three cannot exceed working memory resources if learning is to occur (Paas, Renkl, & Sweller, 2003). Following this assumption, Sweller et al. (1998)

proposed several instructional design techniques based on Cognitive Load Theory. These instructional principles are identified as the goal-free effect, worked example effect, completion problem effect, split-attention effect, modality effects, redundancy effect, and the variability effect.

Goal-Free Effect

The goal free-effect suggests that problems should not be given with an end-goal, because it causes the learner to have to maintain several conditions in working memory while they engage in problem solving. A goal-free problem reduces extraneous cognitive load and aids in schema construction. One example is that a conventional geometry problem will require the learner to find a value for a particular angle, while goal-free problems ask students to find the values of as many angles as they can.

Worked Example Effect

The worked example effect states that providing learners with worked-out examples of problems to study can be just as or even more effective in building schemas and performance transfer than having them work out similar problems themselves. This means that if a multimedia instructional unit was appealing enough to hold the learner's attention and cause the learner to really study the process of a worked-out problem in detail, then it could likely be just as much or more effective than having them work the problem out themselves, at least initially. One strategy that encourages learners to process a worked example at a meaningful, deeper level is **self-explaining**, which we will discuss shortly.

Completion Problem Effect

The key to learning from worked examples, however, is that the examples must be carefully studied, which many learners do not do. Completion problems provide a goal state and a partial solution, and then require the learners to complete the partial solution. This type of problem combines the strong points of worked examples and conventional problems, because the learner must carefully study the partially-worked example and then apply what they have learned to actively solving the problem.

Split-Attention Effect

Split-attention occurs when learners are presented with multiple sources of information that have to be integrated before they can be understood. This principle simply states that instruction should not be designed that causes the learner to have to divide attention between two tasks, such as searching for information to solve a problem or reading a manual while trying to practice a software application on a computer. In the computer example, it is better to have learners read the manual first and then sit down at the computer to practice what they have read.

Modality Effects

This draws from theories such as Baddeley's (1986) theory of visual and auditory working memory subcomponents. It asserts that effective working memory capacity can be increased by using auditory and visual working memory together rather than using one or the other alone. The information that is directed at each channel, however, should be such that it can't be understood in isolation, but needs to be integrated with information in the other channel in order to be fully understood. This of course, is one of the strong points of multimedia instruction, where it is easy to present information visually while also providing related or supporting information through narration, for example.

Redundancy Effect

The redundancy effect occurs when information that can be fully understood in isolation, as either visual or auditory information, is presented to both channels as essentially the same information. Integrating redundant information in both working memories can actually increase cognitive load. What is actually happening when this occurs is a form of split-attention. This strategy can vary, though, depending on the experience of the learner. It is suggested that a diagram with text may be beneficial for novice learners because they need the text to make sense of the diagram, while a similar instructional strategy may become redundant for a more experienced learner and the diagram alone would be more effective. Computer manuals that have minimal text and ample diagrams are another example of a good way to do this. The general message of the redundancy effect is that less is often more when it comes to learning so that cognitive capacity is overtaxed.

Variability Effect

This technique recommends variability of practice because it encourages the learner to develop schemas that aid in transfer of training to similar situations. The more variability in instruction, the more the learner will develop multiple schemas that allow them to recognize common components under different conditions and apply what they have learned to solve problems in other areas.

In addition to understanding working memory and cognitive load for designing multimedia instruction, it is helpful to be familiar with **production system theory**, which seeks to provide a model and explanation for how information is transferred from working memory to long term memory and then retrieved at a later time when it is needed. Some of the production system concepts that we will consider include declarative and procedural knowledge, self-explaining behaviors, and transfer of learning.

A Production System Theory of Knowledge and Learning

Production system theory is important for this discussion in that it further expands the understanding of human working memory and how it interacts with long-term memory to identify goals needed to solve a problem or construct new knowledge. A production system is a model that is based on a set of **condition-action pairs** (if-then statements) known as **production rules** that form the basis of cognitive skills.

For a production to become active or “fire”, it will test incoming information against a pre-determined condition. The stronger the production and the more closely the incoming data meet the condition, the easier it is to trigger the production, which causes a chain reaction, also known as spread activation, which results in a cognitive action of some sort. According to production system theory, learning and automation is actually the process of strengthening these production paths.

One production system that is increasingly being used as a guide for the development of computer-based training (or more specifically, intelligent tutoring systems) is **ACT-R**, which was originally developed by John Anderson for simulating human cognition and understanding how people organize knowledge and produce intelligent behavior.

ACT-R makes several assumptions about how knowledge is represented. The first is that knowledge is stored in two long-term memory structures known as **procedural memory** and **declarative memory**. The second is that a **chunk** represents the basic unit of knowledge in declarative memory, and the third is that **productions** (production rules) form the basic unit of knowledge in procedural memory (Anderson, 1993). One of the most important concepts in ACT-R is this dis-

inction between declarative and procedural knowledge and how the two work together to form human cognition, and that memory and behavior is often a result of some combination or interaction between the two (Anderson & Gluck, 2001). Let's take a closer look at each of them and how they contribute to the acquisition or construction of knowledge.

Declarative Knowledge

Declarative knowledge is factual knowledge that can be reported or described, and its most basic unit is a chunk, which can be hierarchical (chunks within increasingly complicated chunks). Declarative knowledge is spread-activated, and each node has an associated strength, which becomes stronger with use. If these cognitive units are activated frequently, they can eventually become proceduralized and pass to procedural knowledge. The complexity of the chunks that declarative knowledge is able to manage is affected by existing productions in procedural knowledge. This appears to be in line with the discussion of expert/novice knowledge and research suggesting that experts simply have more sophisticated chunks (or schemas) available to them (Chi, Glaser, & Farr, 1988).

In regard to declarative knowledge and the building of schemas, it is important to note that multimedia instruction should not strive to teach with the least amount of cognitive load possible, but at a level that is appropriately tailored to the prior-knowledge of the learner. Research suggests that students with prior knowledge of a subject tend to process the information at a shallower level if the material presented is not challenging, while students with no prior knowledge of the subject do better when cognitive load is kept low (Grace-Martin, 2001). This again raises the point that a multimedia lesson should try to determine the skill level or knowledge of the learner, and then adjust the complexity (cognitive load) to an appropriate level in the learner's Zone of Proximal Development. This can be achieved through a complicated and expensive form of adjusting each step according to the learner's right or wrong answers, or a simple pretest at the beginning of the tutorial, which then suggests certain units.

Procedural Knowledge

Procedural knowledge is dynamic and involves rules, or **productions**, that guide how thinking occurs. Productions are activated through pattern matching and the stronger productions have their conditions met more quickly. For a production to fire, its condition must be matched against information or recognized goals that are in working memory. Although Baddeley (2002) might disagree, Anderson seems to believe that working memory isn't necessarily a separate structure, but rather activated long-term declarative knowledge and temporary structures from encoding processes and productions (Anderson 1983). This cognitive input then follows different paths multiple times as feedback adjusts the weights of connections until the output approximates the expected results. As knowledge is strengthened it is applied at an increasing rate, which eventually results in more capacity being left over (cognitive load is reduced) to acquire new knowledge while the production is being used (Anderson, Boyle, Corbett, & Lewis, 1990). One important conclusion to this is that regardless of one's view about the nature of working memory, Sweller's (1988, 1994) Cognitive Load Theory is relevant.

Acquisition of Knowledge under the ACT-R Model

Declarative knowledge can be acquired quickly from direct encoding of the environment, while procedural knowledge takes longer and must be compiled from declarative knowledge through practice (Anderson, 1993). After a certain amount of practice, the path or production becomes stable and procedural learning has occurred. The conditions under which we learn procedures, therefore, are determined by existing declarative knowledge. Once this happens, it becomes

harder to “rewire” a path, which is why it seems difficult to unlearn something and learn new conflicting information or skills.

Elaboration refers to the process of thinking about and encoding new concepts and prior knowledge together so that the two become more deeply connected. Anderson (1976) states that rich elaboration is critical, because it produces multiple redundant paths for recall in procedural knowledge. Elaboration differs from increasing production rule strength in that strength involves the encoding of a specific memory record, while elaboration creates additional records that can help retrieve the original record (Anderson, 2000).

ACT-R states that knowledge can be acquired either in a passive, receptive mode or in an active, constructive mode. Anderson & Schunn (2000) argue that constructive learning offers no benefits regarding memory and retrieval of knowledge over passive learning other than constructive learning may, at times, provide a redundancy of encoding, but even this generative effect is elusive and not always obtained (Burns, 1992; Hirshman & Bjork, 1988; Slamecka & Graf, 1978; Slamecka & Kkatsaiti, 1987).

Mayer (2004) supports the theory of **constructivist learning** while questioning what he calls the constructivist teaching fallacy, which insists that active learning can only be brought about by active teaching methods such as discovery learning. In other words, the student can be passively sitting in a chair, watching a presentation, but still be very engaged mentally and actively constructing new knowledge. While entertainment and physical activity may be helpful as a change of routine, they do not automatically equate to constructivist learning.

Educators should also be aware that encouraging knowledge construction without structure through activities, such as discovery learning, can have unintended consequences and can lead to the encoding of inaccurate knowledge and incorrect assumptions, while simply presenting material in a meaningful way can be much more efficient and just as effective. All things being equal, it may be best to present the material in a way that allows the learner to construct new knowledge through connecting it with prior knowledge.

There are behaviors that learners can engage in to help them acquire deeper understanding (meaningful learning) through the mental construction of what has been presented. One critical behavior is the practice of self-explanation.

Self-Explanation

Anderson & Schunn (2000) believe that procedural skills are acquired by making references to past problems and then practicing. ACT-R, therefore, is a theory of learning by doing and a theory of learning by example. (Recall the worked example effect in Cognitive Learning Theory.) But simply providing the examples is not enough. A learner must thoroughly understand the examples, and one of the best ways to achieve this is through the activity of **self-explaining** (Alevan & Koedinger, 2002; Anderson & Schunn, 2000).

Chi, Bassok, Lewis, Reimann & Glaser (1989) and Mayer, Dow, and Mayer (2003) demonstrated that students learn better when they apply self-explaining as a metacognitive strategy. Self-explaining is defined by Chi (2000) as the activity of explaining to oneself in the attempt to make sense of new information, usually in the context of learning from an expository text.

Chi is careful to point out that self-explaining is different from talking to or explaining something to others. The focus in self-explaining is simply to understand or make sense of something, while the purpose of talking or explaining to others is to convey information to them. Talking or explaining to others adds the requirement to the learner of monitoring the listener's comprehension, which might prevent the learner from acquiring the knowledge if cognitive load becomes a prob-

lem. It is reasonable to assume that the cognitive capacity that is taxed through talking may hinder the learner from engaging in critical self-explaining behaviors.

An example of how this might be done in multimedia learning would be to intersperse a lesson or activity with breaks where the learner is encouraged to pause for a moment and engage in an activity that cause them to reflect on what has been covered or to re-explain the concept to themselves.

Of course, learning or construction of knowledge is generally only useful if it can be transferred from the situation in which we learn it to another situation where the information or skill needs to be applied.

Transfer

Transfer of knowledge is one of the main goals of learning and instruction. In practice, however, transfer is not as clear-cut as it may first seem. Singley & Anderson (1989) demonstrated that transfer between domains is not usually an all-or-none scenario, but varies depending on how much the two domains use the same knowledge. Task analysis of the knowledge structures is critical for understanding how the knowledge that a learner has acquired in one domain may apply to the other domain.

The importance of this from an educational technology viewpoint is that according to ACT-R, **encoding specificity** becomes critical in the development of tutoring applications and online learning environments if the educational goal is to prepare for performance on a test or task in the short term. The principle of encoding specificity suggests that for recall to occur, the environment in which something is learned should approximate the environment in which it is to be applied. It should not be assumed that transfer of a narrow set of cognitive skills will occur if computer-assisted learning does not closely approximate the actual situation in which the learner is expected to apply the newly-acquired knowledge. When it is said that a person understands a domain in depth, what is really being said is that the person possesses a rich network of readily available declarative chunks and production rules that can be used flexibly to solve problems in many different contexts (Anderson & Schunn, 2000). If transfer is to occur across broad domains, then it should be expected that extensive practice is needed so that a rich network of highly available chunks and productions are developed, which can be used to solve problems flexibly in many contexts.

The obvious application here for multimedia instruction, and where multimedia instruction should excel, is that for transfer to occur, the multimedia learning environment should approximate the situation in which the skill or concept is to be applied as closely as possible, and then have the learner practice as many potential variations as much and as often as possible.

A Cognitive Theory of Multimedia Learning

This now leads us to the topic of how cognitive science can guide us to create more effective computer-based training and multimedia instruction, which Mayer (2001) simply defines as the presentation of material using words and pictures. This definition includes printed materials and emphasizes what Mayer describes as a learner-centered approach rather than the technology-centered approach normally associated with the concept of multimedia. Mayer calls for instruction with multimedia methods that are based on empirical evidence. His Cognitive Theory of Multimedia Learning (Mayer & Moreno, 2002) states that multimedia narration and graphical images produce verbal and visual mental representations, which integrate with prior knowledge to construct new knowledge. According to Mayer and Moreno (1998) and Mayer (2003), the Cognitive Theory of Multimedia Learning is based on several assumptions. First, working memory includes auditory and visual channels, which are equivocal to the phonological loop and the visuo-

spatial sketch pad in Baddeley's (1986) theory of working memory. Second, each subsystem of working memory has a limited capacity, consistent with Cognitive Load Theory (Sweller, 1988, 1994). Third, humans are knowledge-constructing processors who produce meaningful learning when they attend to relevant incoming information, organize the information in coherent representational structures, and then integrate it with other existing knowledge (Mayer, 1996, 1999). Fourth, connections can be made only if corresponding visual and verbal representations are in working memory at the same time, which is similar to Paivio's (1986; Clark and Paivio, 1991) Dual Coding Theory.

Meaningful Learning

Mayer and Moreno (1998, 2003) describe meaningful learning as deep understanding of the material, which includes attending to salient aspects of the presented material, retaining relevant information in both visual working memory and auditory working memory, organizing it into a coherent mental structure, and integrating it with relevant prior knowledge. Mayer (2001) asserts that multimedia learning combining animation with narration generally improves performance on retention tests better than when information is presented as either text or narration alone. More importantly, meaningful learning is demonstrated when the learner can apply what is presented in new situations, and students perform better on problem-solving transfer tests when they learn with words and pictures.

Mayer, Fennell, Farmer, and Campbell (2004) cite evidence that two important ways to promote meaningful learning in e-learning are to design activities that reduce cognitive load, which frees working memory capacity for deep cognitive processing during learning, and to increase the learner's interest, which encourages the learner to use this freed capacity for deep processing during learning. Once again, interest can be stimulated simply by presenting the material in a visually appealing way, accompanied by lively and personable wording or narration. Mayer (2003) lists five cognitive processes that contribute to meaningful learning from multimedia: selecting words, selecting images, organizing words, organizing images, and integrating.

The Science of E-Learning

Mayer (2003) defines a science of e-learning as including three elements: evidence, theory, and applications. According to Mayer, the **element of evidence** means that there is a base of replicated findings from rigorous and appropriate research studies. The **element of theory** requires that there must be a research-based theory of how people learn in electronic learning environments, which yields testable predictions. **Applications** are theory-based principles for how to design electronic learning environments, which themselves can be tested in research studies. As part of his evidence-seeking efforts for the science of e-learning, Mayer (2001, 2003) presents nine major effects which developed out of dozens of studies. These replicated effects are: modality effect, contiguity effect, multimedia effect, personalization effect, coherence effect, redundancy effect, pre-training effect, signaling effect, and the pacing effect. An explanation of each of these nine effects, referred to here as principles (Moreno & Mayer, 2000), follows:

Modality Principle

The modality principle states that better transfer occurs when multimedia combines animation/pictures and narration as opposed to animation/pictures and on-screen text, i.e. students learn better in multimedia messages when words are presented as spoken language rather than printed text. This relates directly to the Theory of Dual Coding which suggests that we have two types of working memory, one verbal and one visual, and that we learn best when both channels are used together, rather than overloading one or the other.

Contiguity Principle

The contiguity principle states that better transfer occurs when corresponding narration and animation are presented simultaneously, both temporally and spatially. Temporal contiguity means that corresponding words and pictures should be presented at the same time, while spatial contiguity means that corresponding words and pictures should be presented near rather than far from each other on a page or screen. In other words, don't place an important visual image on one page or frame, and then discuss it on a preceding or following page/frame without continuing to show the visual image.

Multimedia Principle

The multimedia principle states that better transfer occurs from animation/pictures and narration/words than from words alone. When words and pictures are both presented, learners have the chance to construct verbal and visual cognitive representations and integrate them.

Personalization Principle

The personalization principle states that better transfer occurs when narration is conducted in a conversational style (first or second person) rather than a formal style (third person).

Coherence Principle

The coherence principle states that better transfer occurs when extraneous material such as irrelevant video, animation, pictures, narration, and sounds are excluded. This is where instructional designers who employ gaming technology should be careful. I also like to compare this effect to humorous commercials that we all love and talk about, yet can't remember what the commercial was selling or who the sponsor was.

Redundancy Principle

The redundancy principle states that better transfer occurs when animation and narration are not combined with printed text. When pictures and words are both presented visually, it can overload visual working memory capacity.

Pre-training Principle

The pre-training principle states that better transfer occurs when training on components precedes a narrated animation. If the learner doesn't understand the nature of each component, trying to construct a model of each component while trying to understand how they integrate with each other will quickly overload working memory. It is better to do pre-training on each component so that the learners already possess schemas for them before presenting material that requires the learner to integrate each component into larger schemas. This connects to the concept of chunking and building schemas. Learners have to create low level schemas about a concept, before they can combine them into larger, more complicated schemas.

Signaling Principle

The signaling principle states that better transfer occurs when narrations are signaled. Signaling reduces cognitive load in auditory working memory by providing cues to the learner about how to organize the material. Signaling assists learners in the process of organizing sounds, which can result in deeper, more meaningful learning.

Pacing Principle

The pacing principle states that better transfer occurs when the pace of presentation is controlled by the learner, rather than by the program. Learners vary in the time needed to engage in the cognitive processes of selecting, organizing, and integrating incoming information, so they must have the ability to work at their own pace to slow or pause the presentation if necessary. If the pace of the presented material is too fast, then these cognitive processes may not be properly carried out and learning will suffer.

Putting it all Together

So now that we have covered an extensive range of vocabulary and concepts in just a few short pages, how do we apply all of it? The first thing is to recognize that I have only introduced most of these topics and that extensive literature is available on each of them. Two excellent introductory books for cognitive guidelines on multimedia instruction are *Multi-media Learning* by Richard Mayer (2001) and *e-Learning and the Science of Instruction* by Ruth Colvin Clark and Richard Mayer (2002). Searching for *Intelligent Tutoring Systems* on the Web will also give you an idea of some of the ways that cognitive theory is being applied to computer-based training. The ACT-R site (<http://act-r.psy.cmu.edu>) at Carnegie Mellon is also an interesting place to visit, plus they have an extensive library of freely downloadable publications related to ACT-R.

My own experience in applying cognitive theory to multimedia education has been that my team has been creating online tutorials for several years now. Our mission is to provide training to staff and faculty on computer applications that are supported by the university's Information Technology Services. This includes desktop software, the financial system, and the student-management system. But we are not immune to my own criticism. We have often violated many of the principles discussed in this paper and, in some cases, we continue to violate a few. It is a drawn out process to change a group's practices when they have taken several years to form. The slowness of change can often also be attributed to the new skills or technologies that have to be acquired.

The original search for better solutions came about because I suspected that many of the current educational technology theories were based on not much more than intuition or ivory tower theory that was not really grounded in any scientific method. While they sounded good, I was spending serious money from my budget and I wanted to be convinced, and more importantly I wanted to convince our stakeholders, that our efforts at instructional design were appropriate and really making a difference.

As my team's instructional technology and development skills improved, I began to wonder if we couldn't improve our instructional delivery techniques by basing them in empirically tested theory. It was at this time that we began to look for a set of guidelines to make our instruction more effective and to provide a good return on investment to the university. It was at this point that we discovered and began to implement cognitive theory in our instructional design, specifically the Cognitive Theory of Multimedia Learning (Mayer, 2001, 2003).

The team's e-learning developers and I are currently focused on creating a rapid development process for online Flash-based animated tutorials. To create these tutorials, we first use Macromedia Captivate to capture our instruction on the computer and then we modify them directly in Macromedia Flash MX 2004. We had originally produced most of our videos with live instructor narration. They generally seem to do the job and seemed to follow the principles in the Cognitive Theory of Multimedia Learning. Because live narration was very expensive and time consuming, however, we began to take the shortcut of creating courses which had video instruction, but used captions on the screen rather than narration. From our perspective, this was fine and an effective way to teach our end-users how to use new computer applications. However, once we discovered cognitive theory, the modality effect in particular, it became clear that this was probably not a

very effective way to instruct. According to the modality principle, computer animation with narration produces better transfer than combining animation with on-screen text. This is because you need to try to use both working memory channels (visual and auditory) rather than overloading the visual channel with on-screen text and animation, while leaving the auditory channel practically unused. Some might point out that read text is actually converted into inner speech and held in auditory working memory, but then you could still argue that this is a violation of Sweller et al.'s (1998) split-attention effect which states that you should not design instruction that divides the learner's attention between two tasks: studying the animation and reading the text.

Because we did not have the time or resources to personally narrate every video, we eventually solved this problem by using a high-end computerized voice that transformed the captioned text into narration for each tutorial. This technology allows us to narrate in one or two days what used to take several days or even weeks to accomplish with a human instructor. One additional thing that I would eventually like to investigate is whether there is a difference in performance and transfer between recorded human narration and the slightly unnatural sound of computerized narration.

Adding the computerized narration presented a new problem, however, that we are still working on at the time of this writing. Using Flash, we were able to add a button that toggles the narration on and off so that those who still prefer to read the captions in silence can do so, as well as doing something else such as listening to music or not disturbing their neighbors. But if the user wants to listen to the narration, we currently have not been able to hide the captions in a manner that we deem satisfactory for our end users. Once again, common sense might dictate that it is not that big of a deal to leave the captions up while the narration is playing, but according to the redundancy principle, this can overload the visual working memory and prevent learning. Because of this, our team has vowed to resolve this issue and will have a working solution with 100% performance soon. The end goal will be to have video tutorials in which the learner can toggle both the captions and voice on or off during the playing of the tutorial. It may have occurred to some readers that an apparent solution might be to simply not include the captions, but then the videos would not be accessible to the hearing impaired, which is an important consideration and a violation of Section 508 if you are creating the tutorials for a university or government organization in the United States.

We had also seen other computer-based training where either an animated figure or a video of a person would "lecture" to the learner while text or animation appeared on the screen. My team had considered doing this because we had the capability to do it and it seemed like a good idea. We thought it would add sophistication and another dimension to the tutorials. But the coherence principle recommends avoiding using extraneous material such as irrelevant video or animations, which is what a talking head would be. This is also related to the split-attention effect from Cognitive Load Theory and is something we definitely should avoid. So the team scrapped the idea of including talking figures in our tutorials.

We try to make our narrations as conversational in style as possible, (although this is sometimes difficult with computerized narration) and have tried to avoid using stiff third person narration. This follows the personalization principle which states that better transfer occurs when narration is conducted in a conversational first or second person style.

The pacing principle states that learners should be able to control the speed and pauses in the multimedia lesson. We have accomplished this to a certain extent by including a control bar that Captivate generates when the movie is created. The control bar allows the learner to pause, rewind, and fast-forward the video. One additional feature that our clients have asked for is the ability to speed up or slow down the video, which we still haven't accomplished yet. Controlling the

speed of multimedia is one of the features called for in the pacing principle, however, and we won't be satisfied until we have added that critical feature for our end-users.

The other two principles that I haven't discussed in this section, the pre-training principle and the signaling principle, are critical factors that I also want to include in our instructional planning, but that we currently seem to be doing more by accident than by design. They are equally important, though, and I plan to emphasize them with the team as soon as we have solved some of the other issues mentioned above. I also haven't found a good way to promote self-explaining behaviors in this type of training yet, but it is an area that I plan to investigate as we solve some of the other issues.

Conclusion

Cognitive science provides several empirical theories that provide useful models to suggest ways in which knowledge is constructed and placed in memory. One theory proposed by Baddeley is known as working memory, which temporarily holds and processes information so that it has the opportunity to be stored as representations in long-term memory. The two primary sub-components of working memory are the phonological loop, which processes verbal information, and the visuo-spatial sketch pad, which processes visual information. Cognitive Load Theory builds on Baddeley's model of working memory to propose that since the brain can only attend to and process a limited amount of incoming sensory information, it is important to structure instruction in such a way that working memory is not overloaded. Production system theories such as ACT-R further explain how working memory interacts with prior knowledge to construct new knowledge. Mayer's Cognitive Theory of Multimedia Learning bases his own research on these theoretical foundations to develop a framework that serves to guide the development of effective multimedia instruction.

Because new developments in multimedia technology increasingly have the potential to overwhelm the apparently limited resources of the human brain, it is important that we begin to seriously consider cognitive research and look for ways to apply it more effectively. The exponential growth in computer-based training will precipitate increasing demand for effective learning design in multimedia instruction. Rather than relying on flashy special effects, it is important that instructional designers begin to work within an empirical framework of principles that are driven by the learner, rather than technology. Cognitive science provides a research-based foundation of theories that serve as a grounded starting point for this instruction, as well as further research.

Creating multimedia instruction is gradually becoming easier, but there is still quite a bit to master if it is to be done well. For this reason, our own process of creating multimedia instruction has been a work in progress, and something we expect to continue to refine for years to come. Like everything else in the information technology field, you can never rest or you will quickly find that your skills have become obsolete. When you combine this with the rapidly changing field of cognitive science, it becomes a moving target, and trying to stay current in new technology and cognitive theory presents a formidable challenge. For this reason, my team's attempt to move to an empirical set of guidelines for the creation of our tutorials has been a slow process, but we do know where we want to be eventually. That is, we want to know that we are designing tutorials in a cognitively sound manner, while still producing effective, cost-conscious, and appealing tutorials for our end-users.

References

- Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, 26, 147-179.
- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.) New York: John Wiley & Sons.
- Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, 42, 7-49.
- Anderson, J. R., & Gluck, K. A. (2001). What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr & S. M. Carver (Eds.), *Cognition & Instruction: Twenty-five years of progress* (pp. 227-262). Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 1-34). Mahwah, NJ: Lawrence Erlbaum Associates.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 89-195). New York: Academic Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. D. (1999). *Essentials of human memory*. Hove, England: Psychology Press.
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7(2), 85-97.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Burns, D. J. (1992). The consequences of generation. *Journal of Memory and language*, 31, 615-633.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293-332.
- Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 161-238). Mahwah, NJ: Erlbaum.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.
- Chi, M.T.H., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75). Hillsdale, NJ: Erlbaum.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3, 149-210.
- Clark, R. C., & Mayer, R. E. (2002). *e-learning and the science of instruction*. San Francisco: Jossey-Bass/Pfeiffer.
- Cuban, L. (1996, October 9). Techno-reformers and classroom teachers. *Education Week* 16(6). Retrieved March 20, 2005 from <http://www.edweek.org/ew/articles/1996/10/09/06cuban.h16.html>
- Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.
- De Groot, A. D. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), *Problem Solving*. New York: Wiley.

Cognitive Theory & Multimedia Instruction

- Grace-Martin, M. (2001). How to design educational multimedia: A "loaded" question. *Journal of Educational Multimedia and Hypermedia*, 10(4), 397-409.
- Hirshman, D. L. & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 484-494.
- Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review*, 8, 357-371.
- Mayer, R. E. (1999). Research-based principles for the design of instructional messages: The case of multimedia explanations. *Document Design*, 1, 7-20.
- Mayer, R. E. (2001). *Multi-media learning*. New York: Cambridge University Press.
- Mayer, R. E. (2003). Elements of a science of e-learning. *Journal of Educational Computing Research*, 29(3), 297-313.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14-19.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90(2), 312-320.
- Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12, 107-119.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43-52.
- Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95(4), 806-813.
- Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, 96(2), 389-395.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63, 81-97.
- Moreno, R., & Mayer, R. E. (2000). A learner-centered approach to multimedia explanations: Deriving instructional design principles from cognitive theory. *Interactive Multimedia Electronic Journal of Computer Enhanced Learning*. Retrieved March 22, 2005, from <http://imej.wfu.edu/articles/2000/2/05/index.asp>
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87(2), 319-334.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1-4.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.
- Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 4, 592-604.
- Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language*, 26, 589-607.

- Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. H., Garfield, J. L., & Rissland, E. L. (1995). *Cognitive science: An introduction* (2nd ed.). Cambridge, MA: MIT Press.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257-285.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295-312.
- Sweller, J., Van Merriënboer, J.J.G., & Paas, F.G.W.C (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-29.

Biography



Steve Sorden is the E-Learning & Training Project Manager for Information Technology Services at Northern Arizona University. His current research interests include technology in education, online learning, and technology in developing countries. His 15 years in education include work as a K-12 computer science teacher, high school principal, adjunct faculty, and various positions in higher education information technology. In his “spare time” he is a doctoral student in the Curriculum & Instruction program at Northern Arizona University