8-24-2020

A Multimedia Approach to Improving Retention in Undergraduate Anatomy and Physiology

Kaitlyn Elizabeth Peperone

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

A MULTIMEDIA APPROACH TO IMPROVING RETENTION IN

UNDERGRADUATE ANATOMY AND PHYSIOLOGY

by

Kaitlyn Elizabeth Peperone

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Major: Psychology

The University of Memphis

August 2020

Abstract

This research compared differences in grades among students who utilized an adaptive instructional system (AIS) to study for their undergraduate anatomy and physiology exams. The AIS implemented multimedia techniques by combining machine-generated cloze questions with images drawn from the A&P textbook, specifically chapters 9 (musculature) and 10 (nervous system). The AIS adapted to the performance of each student and optimally chose which question they should see next. Participants studied using this AIS before the exam and were given extra credit for participation. I hypothesized that students who practiced using the AIS will perform better on their exams than those that did not. I hypothesized that multimedia practice will lead to higher scores on questions that directly correspond to the trials in the system. A series of ANOVAs did not show significant results. An exploratory regression revealed that lecture-based questions were easier than lab-based, indicating the importance of the multimedia component.

Table of Contents

Chapter 1

Introduction

Much of a student's success in school relies on their ability to learn from textbooks, whether hardcopies or online versions, as tests, the most common evaluation method in classrooms, are typically derived from textbook content. Issues in content area learning through text arise when a student possesses low reading skills. Low reading skills are a problem for many college students, as only 39% of 12th grade high school students are defined as proficient readers and 69% of high school graduates go onto college (The Nation's Report Card, 2015; US Bureau of Labor Statistics, 2016). This problem may be amplified in content areas that have their own specific set of vocabulary required for comprehending the material, as is often seen in science subjects, such as biology, anatomy, and physiology.

One widely accepted way of improving knowledge and retention in an area is implementing multimedia learning techniques by providing the student with both text and images to learn from (Mayer, 1997). Multimedia learning can take place using technology, like through a computer program, or can be as simple as drawing figures on a chalkboard during a class lecture. The benefits of multimedia learning have been shown especially for those with low prior knowledge in an area. Multimedia instructional techniques can be used as a method of knowledge construction, meaning that a multimedia lesson should be designed to simplify the material into digestible chunks and then scaffold those chunks to promote the construction of a cohesive and complex mental model (Mayer, 2009).

Another way to improve learning from texts is to utilize the testing effect. The idea of the testing effect dates back centuries (Bacon, 1620; James, 1890) and holds that material is better remembered when it is actively tested or recalled. Research showing the validity of the testing

effect implies the value of tests not only as evaluative measures to gauge what a student knows, but also as a method to change and foster knowledge. Testing allows active construction of mental models by challenging the learner to recall information and allowing multiple opportunities to make connections to prior knowledge. Testing effects are versatile and have been proven across temporal differences, subject differences, test format differences, and in combination with multimedia effects (Carpenter, 2012).

Adaptive instructional systems have the capability to combine the effects of both testing and multimedia instruction with adaptive quizzing. Adaptive instructional systems (AIS) are computer systems based off specific individual differences of students and designed to teach a new skill or help a learner acquire new knowledge in a particular field. The idea of adapting instruction to suit individual students is not new, but integrating that concept with computer systems is a modern advance in the field (Reiser, 1987). Adaptivity itself exists on a continuum, and systems often vary on the degree to which they can adapt. AIS can adapt to a wide variety of individual differences, such as specific instructional goals, a student's prior knowledge, or errors that occur during the learning session. The overall flexibility built into AIS allows them to mold to the needs of experiments and classrooms alike.

The proposed research examines the potential of a multimedia adaptive instructional system to improve student grades in an undergraduate anatomy and physiology course. Based on the previous research, this study aims to provide evidence for combining the technology of adaptive instructional systems with the known benefits of the testing effect and multimedia learning to provide a better method of studying for students in a course that is often challenging. By presenting underegraduate students with images with selected cloze sentences in an AIS, MoFaCTS, I expect to see improved accuracy on questions corresponding to those images,

specifically the lab portion of students's anatomy and physiology exams. Improved accuracy is also anticipated as students spend more time practicing within the system.

Learning from Texts

In order to fully understand the problem at hand, we must first look at the prevailing model of human memory. Early cognitive psychologists were highly influenced by computer models, and therefore explained human memory by comparing it to and describing it in terms of such models. The information-processing view of human memory is based off the idea that humans process information they receive, similar to how a computer analyzes information. This view holds that human memory has three main operations – encoding, storage, and retrieval. When information is initially encoded, it is transferred into the memory system. Information is then stored for a variable amount of time before it is retrieved back into awareness for use (Atkinson & Shiffrin, 1968).

In the 1960s, this model of memory was transformed by the proposition of long-term and short-term memory stores. Atkinson and Shiffrin's (1968, 1971) multistore model of memory proposed that information was processed linearly, starting with processing by sensory memory, then temporary storage in short-term memory before being retained, seemingly permanently, in long-term memory. While the idea of short-term memory is nearly universal in theories of memory, it has been increasingly replaced by the concept of a more active working memory. Instead of a passive short-term memory, working memory exists to actively process information and integrate information with prior knowledge before transferring that information into long-term memory (Baddeley, 1992).

Comprehension is a necessary component of the reading process, as readers are tasked with integrating the new information with their prior knowledge. "Good" comprehension of material is achieved when new information is integrated with prior knowledge and that new information is retained (Lipson, 1982). There is a circular causality between comprehension and reading, as individuals read to comprehend but must also comprehend to read efficiently (Olney et al., 2017). Nagy and Scott's (2000) estimate that readers must understand approximately 90% of the written words to understand a text, which also points to this circular relationship between comprehension and reading.

Domain-specific knowledge can greatly improve learning and comprehension from reading texts, especially in academic reading where genres may be more complex (Ricks & Wiley, 2009). Domain-specific knowledge in the form of vocabulary knowledge can effectively expand the limited capacity of working memory (Ericsson & Kintsch, 1995). Typically bogged down cognitive resources are freed up when domain-specific prior knowledge is high, therefore learners can more efficiently attend to the new information being encoded. In academic settings, this ability to allocate cognitive resources towards more complex information by lessening the overall cognitive load through prior knowledge can be quite beneficial if effectively managed.

As previously mentioned, academic vocabulary knowledge can alleviate pressure on working memory, and also acts as a factor differentiating high and low skilled readers. A lack of such vocabulary knowledge often acts as a roadblock to academic success (Nagy & Townsend, 2012). Academic language tends to be inherently denser and more complex than typical English conversational language. This stark difference in combination with the specificity of the vocabulary seen across content areas provides a challenge to readers when learning from academic text. While trying to comprehend the material, readers may fail to participate in

comprehension monitoring simply because they are unaware of the information that is not being processed due to a lack of vocabulary knowledge, or even because they have become disengaged with the complex text (Perfetti & Adlof, 2012).

The difficulty presented by a lack of vocabulary knowledge can be exacerbated in specific subjects. Studies show that results of reading a low cohesion, or more difficult to read, text are highly influenced by prior knowledge in the form of vocabulary knowledge (Voss & Silfies, 1996). In science subjects, like biology or chemistry, learners often do not have any personal experiences or prior knowledge that they can integrate the incoming novel and abstract material with. This lack of prior knowledge in unison with a new and difficult vocabulary can result in poor retention among students (Ozuru et al., 2009).

Early researchers in cognitive sciences used computer models to guide their theories of human memory. Encoding, storage, and retrieval are widely accepted as basic components of models of human memory. As short-term, or working, memory is inherently limited, learning from texts can prove to be a particularly difficult task if a reader's working memory is unsupported by some familiarity with the topic at hand. In order to improve reading comprehension, focus should be shifted towards improving vocabulary knowledge so that the reading process can continue in a smoother manner. It is critical to note that there is no universal best way to improve reading comprehension, but that each individual will have to handle their own specific set of challenges while reading.

Multimedia Learning

Multimedia learning is defined as learning new material from the combination of both words and pictures (Mayer, 1997). Multimedia instruction, or presenting material in both word and picture form, facilitates the process of multimedia learning. Although much of the research

on multimedia learning is relatively recent, the concept itself is not a modern phenomenon, as multimedia learning does not necessitate the use of technology and words and images have been combined for some time to teach concepts (Schnotz, 2014). The goal of multimedia learning is to provide learners with multiple ways to encode the same material in the hopes that it will lead to a more cohesive and unified mental representation. Here I review two main theories behind multimedia learning - Mayer's 1997 Cognitive Theory of Multimedia Learning (CTML) and Schnotz's 2014 Integrative Model of Text and Picture Comprehension (ITPC).

Schnotz's ITPC theory (2014) has less research directly supporting it, and therefore will only be briefly discussed in this paper. The ITPC framework was heavily influenced by Atkinson and Shiffrin's (1968) work on human memory that named three components of memory - the sensory register, short-term memory, and long-term memory. ITPC holds that incoming verbal and pictorial stimuli are processed by specific sensory registers before being held in short-term memory, which has a limited capacity. ITPC also holds that the incoming stimuli play fundamentally different roles. The verbal messages provide descriptive representations, capable of holding more abstract information, while the pictorial messages provide depictive representations, which are powerful in their completeness. It is the job of one's short-term memory to make sense of these representations and to actively combine them into a single representation. Unlike Mayer's Cognitive Theory of Multimedia Learning (1997), Schnotz does not view multimedia effects as a result of the superiority of using dual-coding processes. Rather, Schnotz viewed images and text as inherently different in the functions they serve for human memory.

At its core, Mayer's theory of multimedia learning combines theories of dual-coding, cognitive load, and the generative nature of knowledge (Mayer et al., 1996). Drawing on the

aforementioned theories, Mayer's CTML operates on the assumptions that the human information-processing system is comprised of qualitatively different systems for integrating auditory and visual information, that each system has a limited capacity to store information, and that specific and appropriate processing is required for meaningful retention to take place (Mayer, 2009). This set of three main assumptions has been used to guide instructional delivery methods within multimedia research.

The term "multimedia" can be viewed in several ways, but the most prevailing view is that multimedia refers to the method of delivery of the material. Within the framework of the CTML, this emphasis on the method of delivery of the material is referred to as the presentation-modes view. Information can be presented through verbal modes (e.g., text or verbal narration) or pictorial modes (e.g., static images, animations, or videos). The presentation-modes view is consistent with the idea that learners have different coding systems to process these different types of information, an idea explained by dual-coding theory (Paivio, 1986). While pictures may be able to be described verbally and vice versa, research on dual-coding theory has shown that various modalities are inherently different when it comes to representing knowledge.

From a different perspective, the sensory-modality view holds that "multimedia" means two or more sensory systems of the learner are involved at once (Mayer, 2009). Instead of focusing on the way the material is delivered, this view focuses on the sensory receptors used to receive the information, such as the eyes and ears. For example, a traditional classroom lecture could be considered multimedia instruction since students hear the instructor speak while also viewing images or graphics on a projector screen. More so than the presentation-modes view, the sensory-modality view takes the information-processing activity of the learner into consideration. A crucial facet of this view is that learners process auditory and visual stimuli in

qualitatively different manners. This idea is congruent with the Baddeley's 1999 model of working memory that expands the idea that humans possess different mental channels for processing sounds and visual images. This viewpoint of multimedia provides a solid connection and comparison to Baddeley's model in that the two "halves" of the multimedia instruction can be processed separately and then combined to create a more cohesive representation.

The most predominant idea guiding the development and design of multimedia instruction is that multimedia learning exists to facilitate the construction of new knowledge. Mayer's CTML describes multimedia learning more as a sense-making activity than as a way of adding more information into one's knowledge store. In other words, it is the learner's job to take what is presented to them in various modalities and try to integrate the information into a coherent mental model. Since knowledge construction often varies on an individual basis, this process will occur differently for each learner, but should be guided by the manner in which the material is presented. Since the goal is for multimedia to aid knowledge construction, material should be designed to help the learner decide what the most important ideas are, how to organize the material within their mental models, and how to connect the incoming information with prior knowledge.

As previously mentioned, the dual-channel assumption has much history in cognitive psychology and was popularized by both Paivio (1986, 2006) and Baddeley (1992, 1999). Both theories align with the underlying ideas supporting multimedia learning, but from different viewpoints, as both theorists agree that humans possess separate channels to process visual and auditory information. According to the presentation-modes view of multimedia, verbal material, written or spoken, is processed by one channel, while images, static or animated, are processed by another. This notion is most consistent with Paivio's idea of verbal and nonverbal channels.

In contrast, the sensory-modality view of multimedia focuses on whether the stimuli are initially processed by the eyes or the ears. This view aligns with Baddeley's distinction between the visuo-spatial sketchpad and phonological loop of working memory.

Both previously discussed theories indicate the limited capacity of working or short-term memory, which can hinder learning from multimedia if too much information is presented at once. In order to have ample cognitive resources to integrate or cross-represent stimuli, multimedia instruction must be mindful of this limited capacity of each channel. When images are shown, humans are only able to retain a few images in working memory at a time, and these images are more fragment-like instead of exact replicas of the original image. The same principle holds for verbal stimuli as well. Learners can retain a few words at a time in working memory, and these words typically represent segments of the original verbiage instead of a verbatim replication. This concept of limited capacity has been well studied in cognitive psychology, most popularly by Baddeley (1992, 1999) as previously mentioned, and Sweller (1999, 2005). The limited capacity of working memory points back to the necessity of multimedia learning's guidance of the learner towards the most important parts of the topic at hand.

The final assumption in Mayer's Cognitive Theory of Multimedia Learning is that learners are actively processing the incoming material, as opposed to passively adding disconnected pieces of information into their knowledge base. Learners must attend to the incoming material, organize it, and integrate that information with their prior knowledge – a concept similar to working memory. The goal of this active processing is to create coherent and cohesive mental models that highlight the key components of the presented material, as well as acknowledge the connections between those components. The specific type of active processing appropriate for the content will vary on a case by case basis, but a few examples include making

a chronological list of steps in a process, comparing and contrasting two or more elements, or explaining a cause-and-effect relationship. Active processing will also vary among individuals, but the consistent presence of active processing highlights the importance of multimedia providing a structure for learners to follow in order to minimize extraneous cognitive load.

The integration of words and images is perhaps the most important requirement of multimedia learning. This mechanism involves a change from two separate models (one visual, one verbal) to a single cohesive model in which components from the visual and verbal models are mapped onto one another. This process of integrating mental models occurs when learners are first able to select the most relevant words and images and organize them into appropriate separate mental models before combining them. In addition to combining the visual and verbal mental models, information from long-term memory is also activated and added into the integrated model. This process requires coordination between visual and verbal working memory, as well as long-term memory, and occurs throughout the duration of the multimedia learning session. This critical nature of this process relates to the role of multimedia learning as a sense-making activity, as the learner's goal is to understand the primary structure underlying both the words and images and connect that structure with prior knowledge from long-term memory (Mayer, 2009).

Since multimedia learning requires a large amount of active cognitive processing to appropriately integrate multiple modalities of information, it is important to minimize any extraneous processing or distractions. The coherence principle of multimedia learning holds that leaners will better understand material from a lesson that only contains essential, pertinent material. Although Dewey (1913) argued against adding unnecessary information to learning lessons for the sole sake of making it interesting, more recent research stands in opposition of the

coherence principle and holds that individuals learn better in the presence of seductive text and images that keep them emotionally interested (Weiner, 1990; 1992). According to arousal theory, interesting but irrelevant material keeps learners more engaged, which in turn leads to more learning. Within the framework of multimedia learning, the commonsense approach of arousal theory fails because learners are tasked with making sense of what is presented to them and additional irrelevant material can stand in the way of that through distraction, disruption, or diversion of the learning process (Harp & Mayer, 1998). Research on seductive details shows that adding irrelevant information either reduces or does not improve retention, as students are more likely to remember the interesting details instead of the important ones (Garner, 1992; Garner et al., 1991; Hidi & Anderson, 1992).

If eliminating extraneous and irrelevant information in accordance with the coherence principle is not possible, another solution to decreasing cognitive load is to provide cues that signal the learner towards the most important material. This technique is called signaling. Signaling can manifest in the form of italicizing, bolding, or underlining words, or inserting arrows to accompany images or animations. Signaling does not add any unnecessary information to the material but exists to emphasize the important components that could be lost in the lesson, guiding the learner towards which information to attend to and helping to allocate cognitive resources (Mautone & Mayer, 2001). If multimedia learning is viewed as knowledge-construction, signaling can act as another way to guide the learner towards the most pertinent presented information to further attend to.

Once extraneous processing is controlled through increasing coherence and signaling where necessary, essential processing must be maximized and managed. One way to achieve this is through the segmenting principle, which is an instructional design technique that involves

breaking down a complex idea into smaller "bite-sized" segments. In segmenting, the learner

moves sequentially through the segments of the lesson at their own pace. Since multimedia

learning inherently places a burden on the dual channels of the information processing system,

breaking the material down into chunks can help alleviate this strain. Once material is broken

down into segments, the learner is then better able to integrate the material piece by piece at their

own pace and create a large, cohesive mental model by the end of the session. Segmenting is

likely to have the greatest impact on the learner when the material is elaborate and lengthy and

the learner is inexperienced in that area, therefore it is not always a necessary component of

multimedia instruction (Ayres, 2006).

Together, the framework of both the Integrative Model of Text and Picture

Comprehension and the Cognitive Theory of Multimedia Learning provide solid footing for

research in the area of learning from more than one representation or modality. Research under

both theories has proven the effectiveness of learning in this manner, and the capability of

instructors to provide a better way for their students to learn. While the idea of multimedia

learning is not new or ground-breaking, the years of research going into these two theories add

validity and specificity to what many may already view as a common-sense idea. The

implementation of multimedia learning in classrooms does not have to be a large undertaking for

instructors, as multimedia learning can be lecture, book, or technology-based, and it most often a

combination of all three modalities. The flexibility of multimedia learning lends the method

particularly well to being integrated with the testing effect and being implemented as a more

active type of learning.

The Testing Effect

The concept of testing memory to enhance retention has existed in the philosophy and psychology community for centuries (e.g., Bacon, 1620; James, 1890). The phenomenon known as the testing effect occurs when active practice or testing of material leads to better performance on a later evaluation when compared to re-reading as a study method (Dempster, 1996; Karpicke, 2006; Karpicke & Roediger, 2010; Runquist, 1983; Thompson, Wenger, & Bartling, 1978). While this concept is well-known and widely accepted in psychology research, and some argue it is an underused and underappreciated method in educational practice (Roediger & Karpicke, 2006). Testing has been proven to work effectively across temporal differences, variable test formats, and knowledge domains, pointing at a generalizable effect that adds value to the testing effect as an educational technique (Pashler et al., 2007).

In order to understand the testing effect, it is helpful to compare it to more traditional views of testing. In education, testing often serves as an evaluative measurement used to gauge a student's proficiency in a course (Roediger & Karpicke, 2006). Instructors typically teach course material and then administer a test to determine what a student understands and has retained. In many educational circles, excessive testing is viewed as a roadblock to learning, as it takes potential time and resources away from instruction in the classroom. Tests can have negative effects on students, such as stress and anxiety. The difficulty of using testing as a way to study is often a deterrent, as students are more likely to choose easier methods of studying (Roediger & Karpicke, 2006). Similarly, testing is often not a positive experience for instructors, as most do not like spending the time to create and grade tests (Roediger & Karpicke, 2006). While the negative consequences of frequent evaluative testing can be discouraging, a meta-analysis of thirty-five classroom studies showed that 83% of results indicated positive effects of studying via

13

testing in the classroom setting (Bangert-Drowns et al., 1991). Such results shed light on the benefits of using testing not only as a measurement of knowledge, but also as a way to increase understanding.

While practice via testing in classrooms may not be feasible or easy for instructors, it could provide several opportunities for learners to increase their understanding of material. Frequent testing can encourage students to study more consistently in preparation for those regularly occurring tests, as opposed to students cramming a large amount of material into infrequent study sessions for larger, isolated exams (Fitch et al., 1951). Testing can also encourage students to learn from the feedback they receive by guiding future studying towards questions they answered incorrectly and forcing students to reevaluate their study habits. Testing provides extra opportunities for learners to integrate the incoming information and make sense of it in their own way. These examples of mediated benefits of testing are important as they point towards the practical benefits of testing, however they do not explain the mechanisms behind the testing effect and its efficacy.

Initial research on the testing effect implied that the benefit of testing came from additional exposure to the information (e.g., Thompson et al., 1978). This idea suggests that additional exposure to information leads to overlearning, and that overlearning is responsible for the effects of testing instead of the actual mechanisms of retrieving knowledge. Such explanations for the testing effect arose from studies where effects of testing were confounded with those of total exposure time, thus confounding the results as well. While this explanation seems straightforward and plausible, it does not explain why rereading material outperforms testing in a short-term interval and why testing outperforms studying in the long-term (Roediger & Karpicke, 2006). The idea of overexposure and overlearning also does not explain the ability

of testing to improve transfer, or the recall of information not tested. Such gaps in these theories render them unsatisfactory in accounting for the underlying mechanisms of the testing effect.

As the previously mentioned theory was unable to explain, some component of the retrieval process itself must be a guiding force in the testing effect. Since testing is most often used to improve retention of complex texts, views of reading these complex texts are important to understand. According to Kintsch (1994, 1998), learning from complex text is an active process that requires constructing inferences during and after reading to create a coherent mental representation of the information. If reading itself is a constructive process, it may benefit from other constructive processes that aim to integrate the information. The constructive retrieval hypothesis holds that constructive retrieval of the previously read information may provide opportunities for the learner to recognize or build mental connections or inferences that were not seen during the initial reading of the text (Hinze et al., 2013).

A similar hypothesis, the encoding variability hypothesis, explains benefits of testing as results of the various routes utilized to recall information (McDaniel & Masson, 1985). Each time information is tested, it is reencoded in memory differently due to the varying recall processes. The variety of these encodings equates to an assortment of various retrieval routes available for later recall. The encoding variability hypothesis assumes that retrieval increases the variability of information stored in one's mental representation of a certain concept, and that this increased variability allows for a wider range of cues that would lead to recalling said information.

The idea of transfer-appropriate processing is also useful in understanding the testing effect. Transfer-appropriate processing in the testing effect framework holds that when the processes used for testing memory are the same processes used for encoding the material,

retention will be increased (Morris et al., 1977). Transfer-appropriate processing highlights the important relationship between encoding and retrieval processes, and how this relationship can benefit the learner when those processes are similar. In a more general sense, the transfer-appropriate processing theory serves as a simple explanation of the testing effect that may be beneficial in making a case for more frequent testing in classroom settings.

Bjork and Bjork (1992) developed another theory to explain the testing effect by way of retrieval efforts. This theory distinguished two components of memory - storage strength and retrieval strength. Storage strength refers to the relative permanence of a memory, while retrieval strength refers to the temporary accessibility of a memory trace at any point in time. Bjork and Bjork's model assumes that the two are negative correlated with one another, or that high retrieval strength (e.g., easy accessibility in memory) does not lead to high storage strength in the form of long-term retention. Rather, the theory holds that more difficult and effortful practice, which immediately causes low retrieval strength, is the crucial component of long-term, permanent storage strength. Bjork (1994, 1999) referred to these types of learning in which initial learning is slow and challenging as desirable difficulties, and per this definition, testing qualifies as a desirable difficulty (Roediger & Karpike, 2006). Bjork is not alone in highlighting the importance of challenging a student during learning, as similar ideas have been echoed by theories of the zone of proximal development (Vygotsky, 1978) and appropriate difficulty (Kelley, 1969), as well as the Goldilock's principle (Halpern et al., 2007).

Another potential benefit of testing lies in the appropriate use of feedback through dynamic testing. Standardized tests, like IQ tests or the ACT, provide no feedback during the testing process, and are therefore categorized as static tests. Dynamic tests, on the other hand, generate feedback during testing that acts to assess learning potential and guide future learning

16

(Roediger & Karpicke, 2006). Sternberg and Grigorenko (2001, 2002) have argued the case for dynamic testing as a more efficient and effective way of measuring cognitive skills or knowledge in a domain. In a 2002 study, Sternberg and Grigorenko showed the benefits of dynamic testing and dynamic testing's ability to better measure learning potential when a group of Tanzanian children performed better on a following test only when feedback was provided during the initial test. Dynamic testing enhances the standard testing effect by simultaneously improving learning and evaluating knowledge of a student. The implications of research on dynamic testing point to the opportunity for standardized, static tests to better promote learning potential by adding feedback.

Years of research have proven that testing is a powerful tool to enhance learning, and that the testing effect is a phenomenon that can be generalized across temporal differences, a variety of domains, and various testing formats. Broad theoretical bases, such as transfer-appropriate processing and desirable difficulty, lay a coherent framework for implementing the testing effect in the classroom. While the research side of the testing effect is rich, testing is still an underused concept in most educational settings. Testing can improve student outcomes by raising their grades, but can also function to promote better studying habits, more frequent studying, to lessen test anxiety, and to keep students actively involved in the course material. Overall, the implementation of testing, particularly testing with feedback, has the potential to greatly improve student outcomes and overall learning in school settings. In modern times where technology is rapidly expanding in school systems, creating online systems that allow students study via testing themselves, like adaptive instructional systems, has great potential for success.

Adaptive Instructional Systems

Within formal education, individualized one-on-one tutoring has been heralded as the gold standard (e.g., Bloom, 1984; Cohen et al., 1982). If a teaching technique accounts for individual differences among students and makes changes according to those differences to better suit each student, as is often seen in one-on-one tutoring, it is referred to as adaptive instruction (Como & Snow, 1986). Adaptive Instructional Systems (AIS) are computer systems developed to implement the powerful effects of individualized instruction in order to teach a new skill or help a learner acquire new knowledge in some area. Research in this area exists at the crossroads of computer science, cognitive psychology, and educational research (Nwana, 1990). Such research serves needs of theory development, as well as practical needs in the educational field.

The idea behind adapting content to better suit a student's individual differences is not a new concept and has been viewed as a primary factor in aiding student success since at least the fourth century BC (Como & Snow, 1986). Until the mid-1800s, such individualized adaptation was commonplace in education (Reiser, 1987). Eventually, schools moved towards more group-centered teaching, but researchers in education held their stance that individualized teaching was superior (e.g., Dewey, 1902; Thorndike, 1911). After Cronbach (1957) pointed out that psychology as a field had a duty to examine not only individuals, but also the continuum of variables and interactions that impact said individuals, research has shifted to focus on which specific variables should be examined when building adaptive instructional systems. As AIS have become more sophisticated in nature, many systems now allow for a wide array of variables to be considered as criteria for adaptation, as opposed to the one or two variables that were considered in the early days of development (Park & Lee, 2008).

While the terms adaptive instruction and individualized instruction are often used interchangeably as one in the same, it is important to note that not all individualized instruction is truly adaptive in nature. Most one-on-one instruction can be considered individualized, but it may not be truly adaptive if it is not flexible enough to meet the continuously changing needs of the student. A learning environment is only considered adaptive if it makes changes based on data collected about the student, interactively responds to the student's actions in the environment, and is designed based on common challenges faced by learners in that specific area (Aleven, et al., 2015; Aleven et al., 2013). Any learning environment or AIS may be more or less adaptive than another, as adaptivity is not a binary characteristic and can exist in varying degrees (Aleven, McLaughlin, Glenn, & Koedinger, 2016).

Since the common theme in all parts of the definition of adaptivity is the consideration of data on various learner characteristics, it is critical to decide which factors are the most important. A 1993 account by Jonassen and Grabowski outlined as many as thirty variables to be considered in individualized instruction. The five most predominant variables considered when building adaptive instructional systems are prior knowledge, strategies and errors of the student, affect and motivation, self-regulation, and learning styles (Aleven et al., 2013). Looking at each of these broad categories is beyond the scope of this paper, but it is important to note the main variables implied when talking about individual differences as related to adaptivity.

Three main approaches exist in development of AIS. These approaches are based on the aspects of instruction meant to adapt to differing learners. Macro-adaptive instructional systems typically adapt based off the student's achievement level, general ability, and instructional goals. The aptitude-treatment interaction approach of AIS uses student characteristics, as specified by the system, to determine which instructional strategy would best suit that student. The final

approach, the micro-adaptive approach, continuously diagnoses the student's learning needs and prescribes adaptations to fill those needs throughout the entirety of the learning process. As technology has improved rapidly, micro-adaptive systems have become more feasible to build and therefore more commonplace in the field.

The specific AIS to be used in this study is MoFaCTS, or the Mobile Fact and Concept Training System. MoFaCTS is an micro adaptive quizzing system which proposes student knowledge is composed of chunks, similar to Johnson's 1970 chunk theory. MoFaCTS allows instructors to input whatever material they choose and can present images, sounds, videos, cloze trials, and multiple-choice trials, allowing for great flexibility to suit an instructor's needs. MoFaCTS continuously builds a dynamic model of the student's knowledge and uses that model to drive an algorithm to select the optimal question to present next (Pavlik, Kelly, & Maass, 2017). The system has previously been used to practice Mandarin vocabulary, statistics facts, the human circulatory system, and music intervals, which emphasizes the versatility of the system to mold itself to various content areas.

Chapter 2

Methods

Participants

The participants were recruited from an undergraduate anatomy and physiology course at Southwest Tennessee Community College in Memphis, Tennessee. The student population is roughly 65% African American, 25% Caucasian, 5% Hispanic, 2% Asian, and 3% other. The student population is approximately 62% female and 38% male. We assume that the participants' demographics closely resembled this breakdown, but more specific demographic data was not collected as this study took place in a classroom setting.

Design

This experiment was a repeated measures within subjects design. All participants were exposed to stimuli containing images during their practice for chapters 9 and 10. There were 80 images used for the chapter 9 practice. There were 694 total stimuli, 394 with images and 300 without. There were 20 images used for the chapter 10 practice. There were 308 total stimuli, 98 with images and 210 without.

Materials

The multimedia stimuli were designed using sentences and images chosen from *Hole's Essentials of Human Anatomy & Physiology*, specifically chapters 9 and 10. The cloze sentences were generated by optimal machine techniques previously found to be effective in multiple pilot studies and reflected the most important and pertinent information in each chapter. Chosen images corresponded to the main topic of the cluster of sentences, and the responses were covered in the image by a large blue rectangle. If the image refers to words in the sentence that are not blanked out in that particular trial, the corresponding word was bolded in the sentence in order to reduce confusion. See Figure 1 for an example of the multimedia stimuli used in this study.
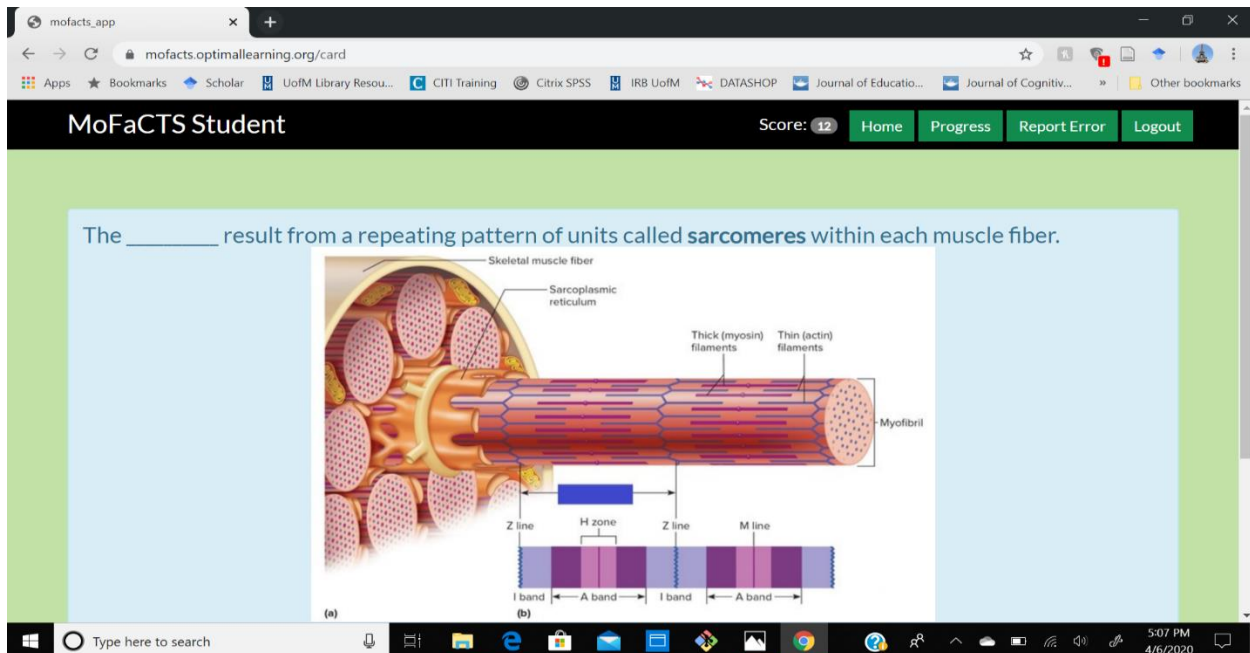
Figure 1. Screenshot of multimedia stimuli used in study.

Procedure

This study occurred as part of a larger project aiming to create an online adaptive study system that generates its own material, but that professors could also tailor to their class as they see fit. Students were encouraged to only use the MoFaCTS practice after attending the corresponding lecture and were offered extra credit for completing one hour of practice. It should be noted that during the Spring 2020 semester, in-person classes were canceled due to the COVID-19 pandemic. Roughly half of the semester was conducted with in-person classes while the other half was conducted online.

After the corresponding lecture occurred, students were provided with a link to the MoFaCTS system. Once they entered the system and clicked on the practice session, students were randomly assigned to see one half or the other of the multimedia stimuli, which were also randomly divided. Students had 30 seconds to respond after being presented with a cloze trial. If they answered correctly, the next trial was shown in 0.75 seconds. If they answered incorrectly,

the correct answer was displayed for 16 seconds before moving onto the next trial. Each student's performance dictated which item they would be presented with next (e.g., concepts answered incorrectly were seen more frequently).

Chapter 3

Results

In order to complete the analyses for this study, a series of steps were taken to transform the data into a more workable format. First, a table of MoFaCTS practice was created, including all stimuli, both with and without images. Next, a concordance table was created. The concordance table contained the anatomy and physiology lecture and lab exam items along with the corresponding cloze trials from the image condition. Another table was created containing each subject's scores (correct or incorrect) on their exam questions. Scores were computed for each subject on the practice items they saw, both with and without images (i.e., correct or incorrect). The scores on the MoFaCTS trials were then compared to the scores on the corresponding exam items.

A repeated measures ANOVA was conducted using the data collected from the chapter 9 trials practiced and exam scores that were transformed into probability values (e.g., count of items answered correctly ÷ count of items seen). There was no statistically significant difference between exam item scores where corresponding MoFaCTS trials were practiced versus exam item scores where corresponding MoFaCTS trials were not practiced as determined by the repeated measures ANOVA ($F(2,13) = .106$, $p = .750$). A one-way repeated measures ANOVA was conducted using the chapter 10 data, again transformed into probability values. A one-way ANOVA was more suitable due to the small amount of complete data available for this chapter's practice. Instead of using the lab and lecture exam scores separately, they were collapsed into

one factor. There was no statistically significant difference between the conditions (items practiced with an image and items practiced without an image) as determined by the one-way ANOVA ($F(5,1) = 1.371$, $p = .586$).

A one-way repeated measures ANCOVA was conducted using the covariates of total trials completed during the MoFaCTS practice and the students's correctness score and comparing. The picture condition probabilities for chapters 9 and 10 were collapsed into one factor, and the same was done for the no picture condition for both chapters. These factors were then compared to the covariates described above. The results were not significant for either total trials ($F(1,1) = 3.436$, $p = .087$) or correctness ($F(1,1) = 1.830$, $p = .199$).

An exploratory mixed effects regression was conducted to examine if students' scores on exam questions could be predicted by their performance on particular questions within the MoFaCTS practice, the chapter the practice belonged to, and whether the MoFaCTS trial corresponded to a lab or a lecture exam question. This model included the random effects of each student, as there were not enough practice items used to examine the effects of the images themselves. The data set used included 324 lines of data for 18 students. Each line contained their anonymized student ID, a MoFaCTS multimedia tria, a coding indicating whether the corresponding exam item was answered correctly (0 or 1), a coding indicating whether the item was seen with or without a picture (0 or 1), the chapter that question belonged to (9 or 10), and lastly whether the question corresponded to an item on the lab exam or the lecture exam of that particular chapter. An interaction was also examined (Item Seen or Not Seen:Lecture) to examine if an item that was seen that corresponded to a lecture exam was predictive of the student's exam score. Table 1 summarizes the results of this analysis.

Table 1

Exploratory Regression with Random Effects of Individual Students

Effect	Estimate	SE	*p*
Fixed Effects			
Intercept	3.16626	1.23047	0.0101
Image Seen or Not Seen	-0.54405	0.79776	0.4953
Chapter	-0.43049	0.74566	0.5637
Lecture Questions	-0.02914	0.73813	0.9685
Image Seen or Not Seen:Lecture	0.57831	0.86282	0.5027

Basic probability descriptives were also run for the entire set of data. Table 2 summarizes the results.

Table 2

Descriptives - Probabilities

	N	Correct	Count	Probability Correct
Image Ch. 9 Lab Exam	8	12	14	.857
No Image Ch. 9 Lab Exam	8	17	20	.850
Image Ch. 9 Lecture Exam	16	68	86	.791
No Image Ch. 9 Lecture Exam	16	65	93	.699
Image Ch. 10 Lab Exam	5	18	22	.818
No Image Ch. 10 Lab Exam	4	10	13	.769
Image Ch. 10 Lecture Exam	6	34	34	1.00
No Image Ch. 10 Lecture Exam	6	33	33	1.00

Chapter 4

Discussion

The testing effect, multimedia learning, issues in learning from texts, and adaptive instructional systems are not new ideas. However, this study took on the challenge of combining all four separate entities into one cohesive system. While the data showed no significant results, the idea behind the MoFaCTS practice remains promising. This study was the first time the capability of integrating images and text was built into the system, as well as deployed in a classroom setting.

As shown in the descriptives table, the total number of participants was quite low ($n = 27$). The maximum number of exam items available for a student to evaluate the results of the training was also low ($n = 21$). These in combination led to a study with low power and no statistical significance. That being said, the purpose of this study was to examine the potential benefits of creating an AIS of this nature, as well as the feasibility of implementing the system in a real class-room setting. From that perspective, this study was a success.

It is crucial to further discuss the implications of the COVID-19 pandemic. Southwest Community College closed its campus for in-person classes after spring break, as did most universities across the country. Students were now expected to manage the load of an already challenging course at home by themselves, along with the other burdens the pandemic undoubtedly placed on them and their families. It would be unreasonable to assume that these new stressors played little to no role in their performance for the rest of the semester. The student population at SWCC is different than typical universities, as a larger portion of students hold full-time jobs or have families to provide for. These responsibilities may have been amplified under the restrictions of the pandemic and their performance in school may have suffered as a

result. Perhaps there was a self-selection bias where students who had more free time on their hands utilized the practice, even though those with little free time may have benefitted the most from the system. Similarly, it is possible that the lack of teacher-student interaction during the duration of remote learning gave students a false sense of security, where they were unaware of how much they did not truly understand.

Results analyzing the covariates as described above do offer some hope. While results were not statistically significant, they were trending in the correct direction, specifically for total trials practiced. These results indicate that perhaps the more trials the student practices, the better their exam scores will be, an idea echoed by the testing effect. It is not far-fetched to think that these results could have crossed into statistical significance if the pool of data had been larger.

Results from the exploratory regression indicate that the lecture exam-based questions in the MoFaCTS practice were easier for students than the lab exam-based items. The lecture-type items were more conceptual or related to functions of anatomy. On the other hand, the lab-type items were more concrete and typically involved labeling a piece of anatomy, or relying on one's capability to identify that anatomy to then answer questions on its function. This result implies that special attention should be given towards lab-type exam questions when building the MoFaCTS system in the future, as these types of questions tend to be more challenging for learners. The fact that lab-type labeling questions were most difficult for students again highlights the need for and importance of a multimedia-based curriculum.

It should be noted that there were limitations of this study. The images used were pulled only from the course text book, but students were tested on images from multiple sources (i.e., lab handbooks, lab practicals, class lectures). The goal of this study was not to examine transfer within the images, and therefore the use of different images was not taken into consideration.

Also, some trials required the student to scroll the browser page in order to see the entirety of the image. This may have confounded our results, leading trials where the image and cloze sentence could be seen without scrolling to have more of an effect than those that were disjointed.

Future work should aim to include expanding the system into different subject areas and grade levels. Future work on a multimedia AIS should remain in classroom settings, as opposed to laboratory settings, in order to obtain the most "real-world" results. A final idea for future research is to create a smoother and more automated process of selecting and integrating images into the multimedia trials. The current process is manual and quite labor-intensive, so the creation of an automated process will only further the user-friendliness of the MoFaCTS AIS.

Conclusions

In conclusion, the overall premise of multimedia learning within an adaptive instructional system seems promising. Years of research point to the idea that the combination of such techniques has the potential to benefit learners. This type of classroom research is crucially important as more technological advances are made in the education field. The COVID-19 pandemic also shined a light on the necessity for and importance of thorough and cohesive remote learning, an area in which the MoFaCTS AIS fits well. Anatomy and physiology lent itself particularly well to fitting the needs of this exploratory study, as students were required to identify muscles and nerve components by sight, as well as explain their function. Research of this sort serves as a critically important junction of theory-based research and real-world applications.

References

Aleven, V., Beal, C. R., & Graesser, A. C. (2013). Introduction to the special issue on advanced learning technologies. Journal of Educational Psychology, 105(4), 929-931.

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2016). Instruction based on adaptive learning technologies. In R. E. Mayer & P. Alexander (Eds.), Handbook of research on learning and instruction (pp. 522-560). New York, NY: Routledge.

Aleven, V., Sewall, J., Popescu, O., Xhakaj, F., Chand, D., Baker, R., ... Gasevic, D. (2015). The beginning of a beautiful friendship?: Intelligent tutoring systems and MOOCs. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), Artificial Intelligence in Education: 17th International Conference, AIED 2015,. (Vol. 9112, pp. 525-528). Berlin: Springer. doi:10.1007/978-3-319-19773-9_53

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*(2), 82–91.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*(3), 287-298.

Bacon, F. (2000). Novumo rganum (L. Jardine & M. Silverthorne, Trans.). Cambridge, England: Cambridge University Press. (Original work published 1620)

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A .D. (1999). *Human memory*. Boston: Allyn & Bacon.

Bangert-Drowns, R.L., Kulik, J.A., & Kulik, C.L.C. (1991). Effects of frequent classroom

testing. *Journal of Educational Research, 85*(2), 89–99.

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies

text from a text-processing perspective: Evidence of improved comprehensibility.

Reading Research Quarterly, 27(3), 251-276.

Beyer, R. (1987). Psychologische Untersuchengen zer Textverabietung unter besonderer

Berucksichtigung des Modells von Kintsch und v. Dijk (1978). *Zeitschrift fur*

Psychologie, Supplement 8.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as

effective as one-to-one tutoring. Educational Researcher, 13, 4–16.

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's model to improve instructional text: Effects

of inference calls on recall and cognitive structures. *Journal of Educational Psychology,*

83(3), 329-345.

Bjork, R.A. (1994). Memory and metamemory considerations in the training of human beings. In

J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–

205). Cambridge, MA: MIT Press.

Bjork, R.A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A.

Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance:*

Interaction of theory and application (pp.435–459). Cambridge, MA: MIT Press.

Bjork, R.A., & Bjork, E.L. (1992). A new theory of disuse and an old theory of stimulus

fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds*.), From learning processes to*

cognitive processes: Essays in honor of William K. Estes (Vol. 2, pp. 35–67). Hillsdale,

NJ: Erlbaum.

Como, L., & Snow, E. R. (1986). Adapting teaching to individual differences among learners. In

    M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed.) (pp. 605-629). New

    York: Macmillan.

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in*

    Psychological Science, 21(5), 279-283.

Cohen, P. A., Kulik, J. A., & Kulik, C. L. C. (1982). Educational outcomes of tutoring: A meta-

    analysis of findings. *American Educational Research Journal, 19*(2), 237-248.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*,

    671–684.

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In

    R. A. Bjork, & E. L. Bjork (Eds.), *Memory* (pp. 317-344). New York: Academic Press.

Dewey, J. (1902). *The child and the curriculum.* Chicago: University of Chicago Press.

Dewey, J. (1913). *Interest and effort in education*. Cambridge, MA: Houghton Mifflin.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review,*

    102(2), 211.

Fitch, M. L., Drucker, A. J., & Norton, J. A. (1951). Frequent testing as a motivating factor in

    large lecture courses. *Journal of Educational Psychology, 42*(1), 1–20.

Garner, R. (1992). Learning from school texts. *Educational Psychologist, 27*(1), 53-63.

Garner, R., Alexander, P., Gillingham, M., Kulikowich, J., & Brown, R. (1991). Interest and

    learning from text. *American Educational Research Journal, 28*(3), 643-659.

Halpern, D. F., Graesser, A., & Hakel, M. (2007). *25 learning principles to guide pedagogy and the design of learning environments.* Washington, DC: American Psychological Society Taskforce on Life-Long Learning at Work and Home.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*(3), 414-434.

Hidi, S., & Anderson, V. (1992). Situational interest and its impact on reading and expository writing. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 213-214). Hillsdale, NJ: Erlbaum.

Hinze, S. R., Wiley, J., & Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *Journal of Memory and Language, 69*(2), 151-164.

James, W. (1890). *The principles of psychology*. New York: Holt.

Johnson, N. F. (1970). The role of chunking and organization in the process of recall. *Psychology of Learning and Motivation, 4*,171-247.

Jonassen, D. H., & Grabowski, B. L. (1993). Handbook of individual differences, learning, and instruction. Hillsdale, NJ: Erlbaum.

Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181-210.

Karpicke, J. D., & Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition, 38*, 116-124.

Kelley, C. R. (1969). What is adaptive training? *Human Factors, 11*(6), 547-556.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 292-303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, MA: Cambridge University Press.

Lipson, M. Y. (1982). Learning new information from text: The role of prior knowledge and reading ability. *Journal of Reading Behavior, 14*(3), 243-261.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*(2), 377-389.

Mayer, R.E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1-9.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: Cambridge University Press.

Mayer, R. E., Mathias, A., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pre-training: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied,* 8(3), 147-154.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43-52.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Topangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*(1), 64-73.

McDaniel, M. A., & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(2), 371–385.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always

    better? Text coherence, background knowledge, and levels of understanding in learning

    from text. *Cognition & Instruction, 14,* 1-43.

Morris, C .D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer-

    appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–

    533.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P. D.

    Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–84). New

    York: Longman.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language

    acquisition. *Reading Research Quarterly, 47*(1), 91-108.

Nwana, H. S. (1990). Intelligent tutoring systems: An overview. *Artificial Intelligence Review, 4,*

    251-277.

Olney, A. M., Pavlik, P. I., & Maass, J. K. (2017). Improving Reading Comprehension with

    Automatically Generated Cloze Item Practice. In E. André, R. Baker, X. Hu, M. M. T.

    Rodrigo, & B. du Boulay (Eds.), *Proceedings of Artificial Intelligence in Education: 18th*

    International Conference (pp. 262-273). Wuhan, China: Springer International

    Publishing.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text

    cohesion in the comprehension of science texts. *Learning and Instruction, 19*(3), 228-

    242.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. Oxford, England: Oxford

    University Press.

Paivio, A. (2006). *Mind and its evolution: A dual-coding approach*. Mahwah, NJ: Erlbaum.

Park, O., & Lee, J. (2008). Adaptive instructional systems. In J. M. Spector, M. D. Merrill, J.

Van Merrienboer, & M. P. Driscoll (Eds.), *Handbook of research on educational*

communications and technology, (pp. 496-484). New York, NY: Taylor and Francis

Group.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J.

(2007). Organizing instruction and study to improve student learning (NCER 2007–

2004). Washington, DC: National Center for Education Research, Institute of Education

Sciences, U.S. Department of Education

Pavlik, P.I., Kelly, C., & Maass, J.K. (2017). The mobile fact and concept training system

(MoFaCTS). *Intelligent Tutoring Systems Lecture Notes in Computer Science,* 247–253.

Perfetti, C., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word

meaning to text meaning. In J. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up:*

Advances in how we assess reading ability, (pp. 3-20). Lanham, MD: Rowman and

Littlefield Education.

Reiser, R. A. (1987). Instructional technology: A history. In R. Gagne (Ed.), Instructional

technology: Foundations (pp. 11-54). Mahwah, NJ: Lawrence Erlbaum Associates.

Ricks, T. R., & Wiley, J. (2009). The influence of domain knowledge on the functional capacity

of working memory. *Journal of Memory and Language, 61*(4), 519-537.

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and

implications for educational practice. *Perspectives on Psychological Science,* 1(3), 181–

210.

Runquist, W. N. (1983). Some effects of remembering on forgetting. *Memory & Cognition, 11*, 641-650.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207-217.

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of of multimedia learning* (pp. 72-103). New York: Cambridge University Press.

Sternberg, R.J., & Grigorenko, E.L. (2001). All testing is dynamic testing. *Issues in Education, 7*(2), 137–170.

Sternberg, R. J. , & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential.* Cambridge, England: Cambridge University Press.

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). New York: Cambridge University Press.

The Nation's Report Card. (2015). NAEP - 2015 Mathematics & Reading at Grade 12 - Reading - National Average Scores

Thompson, C. P., Wenger, S. K., & Bartling, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning & Memory, 4*(3), 210-221.

Thorndike, E. L. (1911). Individuality. Boston, MA: Houghton Mifflin.

US Bureau of Labor Statistics. (2016). College Enrollment and Work Activity of High School

      Graduates News Release.

Voss, J., & Silfies, L. (1996). Learning from history text: the interaction of knowledge and

      comprehension skill with text structure. *Cognition and Instruction, 14*, 45-68.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process.* Cambridge,

      MA: Harvard University Press.

Weiner, B. (1990). History of motivational research in education. *Journal of Educational*

      Psychology, 82(4), 616-622.

Weiner, B. (1992). Motivation. In M. Alkin (Ed.), *Encyclopedia of education research* (6th ed.;

      pp. 860-865). New York: Macmillan.