TWO STAGE PROCESS MODEL OF LEARNING FROM MULTIMEDIA: GUIDELINES FOR DESIGN

A Dissertation
Presented to
The Academic Faculty

By

Jesse S. Zolna

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Psychology

Georgia Institute of Technology

April, 2008

TWO STAGE PROCESS MODEL OF LEARNING FROM MULTIMEDIA: GUIDELINES FOR DESIGN

Approved by:

Dr. Richard Catrambone, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Paul Corballis
School of Psychology
Georgia Institute of Technology

Dr. James Foley
College of Computing
Georgia Institute of Technology

Dr. Mark Guzdial
College of Computing
Georgia Institute of Technology

Dr. Wendy Rogers
School of Psychology
Georgia Institute of Technology

Date Approved: March 26, 2008

I would like to dedicate this -- the most thorough work of my life -- to my Mother, whose example, support, and encouragement was important to help me complete this project. The way she handled herself over the past several years made me proud. Though she is not here to express it, I am sure she would feel the same to know that I have completed this project. I love and miss you mom, and I will remember you every time I see those three little letters after my name.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Richard Catrambone and the rest of the Problem Solving and Educational Technologies Lab for their advice, assistance, and friendship. Without them I could not have come this far. I would also like to thank all of the members of my committee, Paul Corballis, Mark Guzdial, Jim Foley, and especially Wendy Rogers for their input, suggestions, and help. I would like to thank the American Psychological Association who supported this research with a Dissertation Award and the people at the Center for Assistive Technology and Environmental Access who employed me during my time at Georgia Tech and who were very generous and flexible in allowing me to complete this project. Last, but not least, I would also like to acknowledge the support of my father, Fred Zolna, and his father, the original Dr. Zolna (Martin), both of whom did all they could to encourage me to pursue knowledge as an end all its own.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Using multiple media to create instructional materials can increase the likelihood that people will learn from a lesson (Ainsworth, 1999; Mayer, 2001). Theories of learning from multimedia suggest that when media include two modal forms (e.g., visual and auditory), learning is improved by activating modally segregated working memory subsystems, thereby expanding the total cognitive resource available for learning (Mayer, 2001; Sweller, 1999). However, a recent meta-analysis suggests that the typical modality effect (use of narrations and diagrams [i.e., multimodal] leads to better learning than use of text and diagrams [i.e., unimodal]) might be limited to situations in which presentations are matched to the time it takes for the narration to play (Ginns, 2005). This caveat can be accounted for by the differences in ways that people process unimodal and multimodal information, but not by the expansion of working memory explanation for modality effects (Tabbers, 2002). In this paper, I propose a framework for conceptualizing how people interact with multimedia instructional materials. According to this approach, learning from multimedia requires (1) creating mental codes to represent to-be-learned information and (2) forming a network of associations among these mental codes to characterize how this information is related. The present research confirms, in two between-subjects experiments, predictions from this model when presentation pace and verbal presentation modality are manipulated to accompany static (Experiment 1) and animated (Experiment 2) diagrams. That is, the data suggest that learning from unimodal presentations improved as presentation pace was slowed, whereas learning from multimodal presentations did not change as presentation pace was slowed. A third experiment also confirmed predicted patterns of eye movement behavior, demonstrating patterns of increasing dwell time on pictures and switches between media as pace was slowed for unimodal presentations but not multimodal presentations. It is concluded that the parallel patterns of learning outcomes and eye-movement behavior support the proposed model and are not predicted by other models of learning from multimedia instructions. This improvement in predictions of the effects of manipulating design elements (e.g., presentation pace and verbal presentation modality) on learning can help designers as they consider what combination of resources (e.g., classroom time or equipment for multimodal presentation) to devote to instructional design.

CHAPTER 1: TWO STAGE PROCESS MODEL OF LEARNING FROM MULTIMEDIA: GUIDELINES FOR DESIGN

Multimedia presentations include two or more media such as text, graphics, images, audio, or video. Leveraging technology to create instructional materials using more than one medium has a number of learning advantages compared to more traditional instructional materials that use only one medium (see Eskicioglu & Kopec, 2003), including increasing the likelihood that people will learn the information included in a lesson (Ainsworth & Fleming, 2006; Mayer, 1989).

It is possible that the addition of one medium to another improves learning because the flexibility in using multiple media types enables designers to choose the appropriate medium to express complementary information in computationally efficient ways (Ainsworth, 1999; Ainsworth & VanLabeke, 2004). Larkin and Simon (1987) demonstrated that diagrammatic materials are better for making relationships between physics or geometry concepts explicit, and symbolic materials (e.g., words) are better suited for expressing and emphasizing precise details. These types of media can have differential effects on learning, even when they contain equivalent information (Schnotz & Bannert, 2003). Some instructional materials require more than one medium to adequately present all of the necessary to-be-learned information (Bieger & Glock, 1986).

Combining diagrams or pictures with complementary words has repeatedly been shown to further improve learning outcomes with various tasks and measures. For example, compared to conditions using either text or pictures only (each of which contained all of the information necessary to the lesson) to teach procedural instructions, the use of text and pictures together led to better recall, verification of the order of steps, and reduced time to verify the order of steps (Brunyé, Taylor, Rapp, & Spiro, 2006). In a another series of studies adding visual illustrations to a verbal-descriptive summary used to teach students how lightning is formed improved fact recall and transfer (Mayer, Bove, Bryman, Mars, & Tapangco, 1996) and adding visual illustrations to a verbal-description of how brakes work improved conceptual recall and problem solving (Mayer, 1989; Mayer & Gallini, 1990). In addition, adding visual screen shots to a text instruction manual led to the formation of a stronger mental model of a complex software application, improved identification of on screen window elements and objects, and speeded users' ability to locate window elements and objects (Gellevij, Van Der Meij, De Jong, & Pieters, 2002).

However, combinations of media need to be carefully configured to aid learning (Scanlon, 1998). On one hand, recall and transfer following learning from text books can be improved by replacing the visual text with equivalent audio-narrations (Mayer & Moreno, 1998; Moreno & Mayer, 1999; Mousavi, Lowe, & Sweller, 1995) or by simply integrating text and pictures as opposed to keeping them separate (Chandler & Sweller, 1992). On the other hand, many studies have also shown that learners can fail to benefit from the use of multimedia instructional materials (Ainsworth, 1999; Mayer & Anderson, 1992) when the text is very rich on its own (Mayer, 2001), or when limited time is provided to study materials (Tabbers, 2002).

The goal of this study is to explore how to configure multimedia instructional materials to take full advantage of being able to choose among presentation alternatives while avoiding making changes that do not improve learning. To do so, the present research offers a model to explain and predict the effects of multimedia configuration on learning. This model will help answer the question of how outward changes made to the presentation of content-equivalent instructional materials can influence learning. In the following sections, I will define multimedia instructional materials, present a two stage process model of how people learn from multimedia instructional materials, and explore how this model relates to other views of how people learn from multimedia instructional materials. Finally, I will report two experiments performed to examine learning from different configurations of multimedia instructional materials. This study will serve to compare predictions of the proposed model with other models that are frequently cited in the literature. A third experiment will measure eye-movement to examine the differences in how learners observe multimedia instructional materials and link these behaviors to different learning outcomes observed in the first two experiments. In summary, I will use this model to suggest improved guidelines for the design of multimedia instructional materials.

1.1 What are multimedia instructional materials?

Much of the research involving multimedia instructional materials defines them as a "presentation involving words and pictures that is intended to foster learning" (Mayer, 2001, p. 3). A broader definition of multimedia instructional materials might be "any presentation combining more than one format, whether it is within a single sensory modality … or across modalities" (Brunyé et al., 2006, p. 918). Multimedia instructional materials have taken many forms in design research, including movies (Baggett, 1989), graphs and equations (Kozma & Russell, 1997), text instructions accompanying diagrams

(Sweller, 1999), or animations and narrations (Mayer, 2001). In addition, multimedia instructional materials have been designed to teach a wide range of topics including scientific cause and effect systems (Baggett, 1989; Butcher, 2006; Mayer, 2001; Sweller, 1999), computer algorithms (Byrne, Catrambone, & Stasko, 1999), chemistry formulas (Stasko, Catrambone, Guzdial, & McDonald, 2000) and cognitive skills (Tabbers, Martens, & van Merrienboer, 2004). Establishing an understanding of the characteristics of multimedia instructional materials is important to being able to conceptualize how configurations can be manipulated and how learning can be measured.

Central to the concept of *multimedia* instructional materials is the combination of at least two media. A *medium* (the singular of media) is the mode of rendering information that arises from interaction between a presentation modality and a symbol system (Salomon, 1994). Media can differ by modality in which they are presented or the means by which they express information content. An example of media that differ by modality might be verbal-*text* explaining the correlation between two variables versus verbal-*narration* explaining the correlation between two variables. An example of media that differ by the means (i.e., the symbol system) by which they express information content might be verbal-text *explaining* the correlation between two variables versus a Venn diagram *showing* the correlation between two variables. In addition, media can differ along both modality and means, for example verbal-narration explaining the correlation between two variables versus a Venn diagram showing the correlation between two variables. See Figure 1 for an example of alternative presentations of equivalent content.

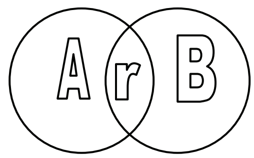
Presentation modality Presentation means	Visual presentation	Auditory presentation
Verbal content	verbal-text: Variable A and Variable B have a correlation of r	verbal-narration: *"Variable A and Variable B have a correlation of r"*
Spatial content	Venn diagram: 	Sonification: *Two similar sounds that suggest the two items are similar.*

Figure 1: Various alternative media that can be used to convey the fact that two variables are correlated.

A large amount of research has been conducted to examine the educational effects of manipulating the media used to present verbal and visual content (e.g., Mayer, 2001; Sweller, 1999). The present study will manipulate the modality of verbal materials, creating content equivalent text passages and narrations. There has been quite a bit of research performing similar manipulations (see Ginns, 2005 for a review). However, there remains a debate over the root of learning effects attributed to manipulation of verbal presentation modality (c.f., Tabbers, 2002). Moreover, there are at least a few caveats for design guidelines related to choosing the presentation modality of verbal materials (Mayer, 2005a; Sweller, 1999; Wickens & Hollands, 2000). The goal of the present study is to hold the means of presentation constant in order to isolate the effects of manipulating presentation modality. The hope is to improve our understanding of the underlying mechanisms of how people learn from multimedia.

Another central component to the concept of multimedia *instructional materials* is their intent to help observers learn discrete facts and the basic knowledge structure that comprises a lesson (Chambliss & Calfee, 1998; Cook & Mayer, 1988). For instance, a multimedia lesson explaining a scientific cause-and-effect process might present a variety of states as well as a system in which changes in one state cause predictable changes in another state (e.g., Mayer & Chandler, 2001). Alternatively, a multimedia lesson presenting foreign language word pairs might present a list of words and indicate how they should be matched into dyadic pairs (e.g., Dubois & Vial, 2000). The types of facts and the basic knowledge structure of a lesson have implications for learning goals and measurement. To link instructional design and behavioral outcomes Gagne (1972) defined five task independent learning domains (see Table 1) that have different requirements for learning and lead to different effects on behavior, such as test performance outcomes. According to Gagne, these learning domains are categories within which generalizations about how people learn can be drawn. These categories can be used to help instructional designers define and examine learning performance goals independent of specific course content (Gagne, Briggs, & Wager, 2005).

Table 1: Categories of learning based on Gagne, 1984.

Domain	Characteristics of learning outcome
Intellectual skills	Learning of concepts, rules, and procedures. Shown when a person is able to apply a sequence of concepts representing condition and action, e.g., understanding the combination of forces that cause lightning and identifying the key elements of a situation which is preventing lightning from occurring.
Verbal information	Declarative knowledge. Stating previously learned materials such

	as facts, concepts, and principles, e.g., listing the four chambers of the heart.
Cognitive strategies	Problem solving. Employing personal ways to guide learning, thinking, acting, and feeling, e.g., devising a corporate plan to improve customer relations.
Motor skills	When gradual improvements can be attained by repetition. Executing performances involving the use of muscles, e.g., doing a triple somersault dive off the high board.
Attitudes	Internal states inferred from behavior. Choosing personal actions based on internal states of understanding and feeling, e.g., deciding to exercise daily as a part of preventive health care.

For the present research, I have chosen to examine multimedia instructions using verbal and visual materials to help learners understand a scientific cause-and-effect system. Even though the content of a scientific cause-and-effect system can be quite complex, a complete lesson can be taught in a relatively short learning session, and often people can learn the content quite well. Perhaps for these reasons, learning from this type of lesson has been examined extensively in past (Baggett, 1989; Butcher, 2006; Mayer, 2001; Sweller, 1999). This fact has the additional advantage of being able to relate the research proposed here to a large body of extant research. Moreover, multimedia instructional materials intending to teach a scientific cause-and-effect system have easily identifiable and measurable learning goals. A *verbal information* learning outcome might be seen as identifying and committing the to-be-learned information to memory in the form of verbal labels and individual concepts. This forms the knowledge structure upon which *intellectual skills* depend (Gagne et al., 2005). In the case of scientific cause-and-effect systems, intellectual skills might be seen as developing an understanding of the rules and principles inherent to the relationships between concepts or system states.

1.2 How do people learn from multimedia instructional materials?

Learning from any instructional materials depends upon the acquisition or reorganization of the cognitive structures humans use to store and process information (Good & Brophy, 1990). That is, to learn the information contained in a multimedia lesson, learners must perceive information presented in a variety of ways; the most common presentations of multimedia lessons target the visual sense via text or pictures and/or the auditory sense via narrations. Learners must then translate the to-be-learned information (regardless of presentation type) into an internal symbol system, elaborate on those internal representations, and associate them with other relevant information (Salomon, 1994). For instance, to extract information about shape from a diagram, learners must be able to examine each side of the diagram or to extract information from a sentence they must examine each word. Examining only one side or just a few words would not seem to be sufficient.

According to Anderson's Human Associative Memory (HAM) (1973), people store external information (e.g., to-be-learned information perceived from instructional materials) in the form of a binary graph structure consisting of *nodes* that represent ideas, and *links* that represent relations or associations between these nodes. See Figure 2 for an illustration of a simple diagram and a sample tree structure representing how the diagram might be stored in memory according to HAM. In this type of memory structure, idea nodes are abstract representations of facts and concepts that derive additional meaning (e.g., specific details) from associations with other abstract idea nodes (Anderson & Bower, 1973). From this encoding scheme emerges semantic information such as the meanings represented by the words and the implied relationships between them in Figure 2b.

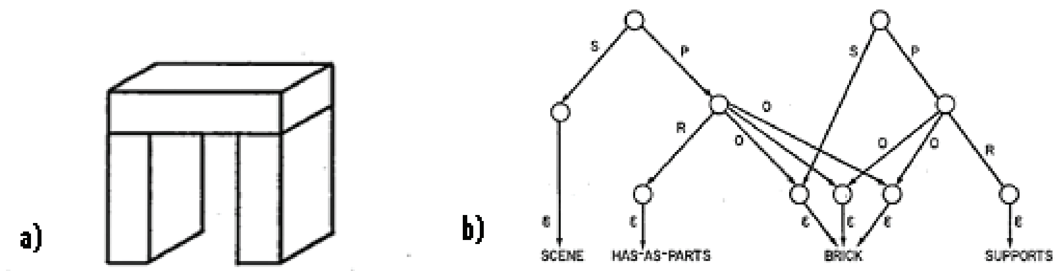


Figure 2: Forming mental codes to represent to-be-learned information (1) and identifying their associations (2) to form a coherent mental model in the two stage process model of learning from multimedia.

Following from HAM, it could be that to-be-learned concepts presented via multimedia instructional materials might be represented in memory via assemblies of nodes and their associations. These cognitive structures, which might be conceptualized as mental codes (e.g., Baddeley, 2003; Barnard, 1999; Lecerf & de Ribaupierre, 2005) are a conversion of sensory input into a unique combination of nodes representing abstract properties and the associations that give them meaning. This concept of mental codes has been theoretically tied to active thought in working memory (Baddeley & Hitch, 1994) and the storage of information in long term memory (Anderson, 1995), as well as complex cognitive activity such as learning (Miyake & Shah, 1999). Moreover, it seems plausible that the temporary working memory storage of mental codes representing to-be-learned information is integral to learning from instructional materials (Gagne et al., 2005).

Multi-component working memory approaches to cognitive architecture are based on the premise that there might be some limit to the amount of information that can be active in working memory at one time (Miller, 1956) and these cognitive resource limitations can be traced to modality based subsystems with individual capacity limitations (e.g., Baddeley, 2003; Barnard, 1999; Paivio, 1986; Penney, 1989; Wickens & Liu, 1988). A mental code representing information from one particular medium might include a node with properties related to the sensory system in which the information was perceived or contain an element that imparts visual or phonological properties to a memory structure (Baddeley & Hitch, 1994), or include information that is inherently visual-spatial or verbal-symbolic (e.g., Paivio, 1986; Wickens & Liu, 1988). Therefore, these modality based subsystems are assumed to be dependent upon perception.

Framing working memory around multiple independent subsystems and their individual capacity limitations enables the alignment of theoretical predictions with empirical data. For instance, research in multimedia learning has shown that people better recall a list of words when corresponding pictures have been presented with each word (Paivio, 1986) or when the list was presented in auditory form (Penney, 1975). The assumption that sensory (modal) aspects are inherent to mental codes grew out of the hypothesis that the use of multiple media to represent the same information improves recall by creating redundant encoding of the semantic information (Paivio, 1986; Penney, 1980). That is, two separate mental codes might represent the same semantic information but differ based on some modal property. Either mental code (that differ based on sensory modality but represent the same word) might lead to recall when prompted, improving verbal learning outcomes, such as recall of word lists (Gagne, 1972; Gagne et al., 2005). That is, providing the learner with two alternative mental codes by which to remember a word in a list might increase that chances that the word is accurately recalled.

Unfortunately, the story might be different for instructional material; empirical research suggests that simply adding media, even of a new modality, containing redundant information to more complex materials (e.g., instructional materials describing a scientific cause-and-effect process) does not improve learning (Kalyuga, Chandler, & Sweller, 1999; Mayer, Heiser, & Lonn, 2001; Mousavi et al., 1995). This might be because, though mental codes have the potential to contain information related to perceptual characteristics, they are not fixed to their perceptual form (Scaife & Rogers, 1996) and might be translated into the same code or a different code than the instructional designers intend, thus affecting learning in unintended ways. As learners observe a lesson, conscious thought aimed at encoding to-be-learned information into mental codes can restructure and edit their meaning (Anderson & Bower, 1973) or partially enhance information (Cowan, 1988) through connections made to memory (Penney, 1980), related semantic information (Barnard, 1999), or complementary information in other memory stores (Baddeley & Hitch, 1994). In this way, verbal information can evoke mental codes containing visual-spatial information (Baddeley, 2003) to the extent that the words convey visual-spatial information (De Beni, Pazzaglia, Gyselinck, & Meneghetti, 2005). Therefore, after information is perceived (i.e., converted from media into a mental code), modal properties of a mental code might not match the modal properties of its presentation (Schnotz, Bannert, & Seufert, 2002). Moreover, there is evidence that a given task may be performed using either a verbal or spatial strategy regardless of the physical presentation of relevant information (Wickens & Liu, 1988). For example, in determining appropriate headings in air intercept control, the controller may adopt an arithmetic (verbal) mode of computing relative bearings and headings or a spatial visualization mode of imagining a triangle (Weinstein, 1987). It is also possible that identical mental codes are formed to represent a word whether it is presented via an auditory or visual medium (Penney, 1989). This could mean that a to-be-learned concept is remembered via one single mental code even when it is presented redundantly via text, speech, and or pictures.

If it is possible that redundant but separate (and even modally distinct) media evoke a single mental code, improvements in recall related to adding content equivalent pictures to words might not solely be the result of the formation of additional mental codes, but instead the formation of associations *between* mental codes (Clark & Paivio, 1991). Learning theorists often suggest that learning is improved by forming associative and referential connections between mental codes (Barnard, 1999; Kieras, Meyer, Mueller, & Seymour, 1999; Paivio, 1986) and that learning might be aided by developing a comprehensive network of associations among mental codes representing to-be-learned information (Anderson, 1995). That is, in the case of word lists, a picture might lead to an association with a similar visual representation of an item in the list or auditory presentation might lead to an association with a similar sounding word in memory. In the case of multimedia instructional materials conveying scientific cause-and-effect processes, if the visual information conveys context and the verbal information conveys details (Larkin & Simon, 1987), learners must be able to connect the context to the details for a complete understanding of the lesson. This possibility is supported by research suggesting that the ability to learn the relationship between points on a map seems to depend upon encoding a representation of multiple subsections of the map and their relations to each other (Zimmer, 2004) and that logical reasoning in word problems might be based on forming a mental model representing actors in the problem and the relations between them (Goodwin & Johnson-Laird, 2005).

In fact, research suggests that encouraging the development of associations between complementary verbal and visual materials might improve learning from multimedia instructional materials (Good & Brophy, 1990; Mayer & Sims, 1994; Schnotz & Bannert, 2003). Improvements in learning from multimedia instructional materials are thought to be related to encouraging learners to integrate complementary words and pictures into an interconnected network of ideas (e.g., Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Brunyé et al., 2006; Craik & Lockhart, 1972; Mayer, 1989; Penney, 1980). In addition, a recent study used a think aloud protocol to reveal that the primary learning benefit related to adding diagrams to text instructions was that diagrams support the generation of inferences based on integrating information between media (Butcher, 2006).

In the case of a scientific cause-and-effects lesson, learning intellectual skills might depend upon by the ability to identify the ways that words and pictures are associated, combining corresponding representative structures in ways that can produce novel associations (see Gagne, 1972; Gagne et al., 2005). That is, associating two separate mental codes might lead to the ability for learners to predict the relationship between elements of these mental codes. It seems that learners do this automatically when observing multimedia instructional materials. Eye movement research on the integration of pictures and text has shown that in most cases learners first read (at least part of) the text and then switch to the picture to integrate the verbal and the pictorial information (Hegarty, Mayer, & Green, 1992; Rayner, Rotello, Stewart, Keir, & Duffy, 2001) and that as people listen to a story or follow instructions, they quickly move their eyes to those elements in an array that are most closely related to the words currently heard (Cooper, 1974).

Based on this understanding of human cognitive architecture, learning a scientific cause-and-effect process might be considered a two stage process (c.f., Mayer & Chandler, 2001). The first process requires creating mental codes to represent relevant concepts and other to-be-learned information that are presented in a lesson. The second process requires identifying relevant associations among these mental codes to characterize how the information contained in complementary materials is related. See Figure 3for an illustration representing the two stage process model of learning from multimedia instructional materials.

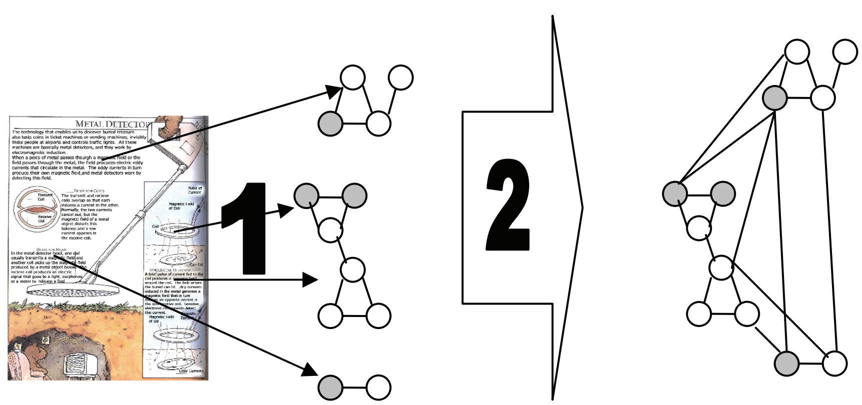


Figure 3: Proposed two stage process model of learning from multimedia: stage 1 is the perception of information and formation of mental codes, stage 2 is the formation of associations between mental codes aided by the reexamining materials.

1.3 Measuring learning behavior

The two processes of learning outlined above might also be separable in terms of how they influence learning: the formation of mental codes might lead to verbal learning outcomes, but the ability to combine these structures in ways that form new associations might lead to developing intellectual skills (see Gagne, 1972; Gagne et al., 2005). That is, the configuration of multimedia instructional materials might improve learning by facilitating the acquisition of mental codes (which should lead to improvement in verbal learning outcomes such as fact recall) and the formation of associations among metal codes (which should lead to an improvement in developing intellectual skills – as defined by Gange - outcomes such as knowledge transfer).

Ideas, concepts, and basic information pertinent to a multimedia lesson can be presented to learners via various media (e.g., narration, text, and pictures). The proposed model suggests that future recall of this factual information requires the formation of a mental code representing each of to-be-learned concepts. To form mental codes representing facts contained in a lesson, learners must first perceive that information, whether it is presented visually or auditory. The formation of these codes leads to verbal information learning outcomes such as the capability to consistently label an object or object class from the lesson; after learning, observers should be able to label concepts (Gagne, 1972). Therefore, verbal learning outcomes can be measured by observers' ability to recall simple facts and recognize elements from the lesson that requires understanding simple concepts (Mayer, 2001).

Once facts and their basic relationships can be identified, learners must form a coherent mental model that helps them understand how facts are interrelated. In addition, this sometimes leads to the development of novel information. According to the two stage process model of learning from multimedia instructional materials, the development of intellectual skills requires a comprehensive network of associations among mental codes in memory. This network of associations might be more easily formed when presentation allows the learner to determine how materials are related and identify the correspondence between distinct media. The learning of these rules and productions leads to intellectual skill learning outcomes; in the case of a scientific cause-and-effect system, learners should acquire the capability to determine rules and principles inherent to the scientific system, understand relationships among parts of the system, and how changes in parts of the system affect other parts of the system (Gagne et al., 2005). Therefore, intellectual skill outcomes can be measured by observers' ability to perform knowledge transfer to novel situations including redesign, troubleshooting, and making predictions of what happens as a result of making changes in the system.

The two processes of learning outlined above might be separable in terms of how learners interact with instructional materials to complete each process. It might be that the formation of mental codes occurs relatively effortlessly at perception, but stable memory modification requires effortful processing of to-be-learned information (Fisk & Schneider, 1984). Therefore, the formation of mental codes might require time to perceive to-be-learned information, while the formation of associations among codes might require time for deliberate searching of the materials to determine how information is related. Each of these processes requires that learners have enough time to complete them. The capability to complete these processes might thereby be observed separately in eye movement behavior. The formation of mental codes might be operationalized as time viewing or hearing the materials. That is, measuring how long learners look at materials might be one way to measure the hypothetical construct of forming mental codes of the information contained in those materials. Forming associations might be operationalized as looking back and forth between related materials or viewing materials related to currently playing bits of auditory presentation. That is, measuring patterns in the order of viewing complementary materials might be one way to measure this hypothetical construct of identifying and forming associations between complementary mental codes derived from those materials. Understanding these processes and how design influences them might help instructional designers better understand the effects of design decisions on learning.

1.4 Understanding how to configure multimedia instructions to foster learning

Currently, there are two major theoretical approaches to using an understanding of human cognitive architecture to examine how the configuration of multimedia instructional materials can influence learning outcomes (e.g., test performance) (Mayer, 2001; Sweller, 1999). Subsequent research has led to related guidelines for how to configure multimedia instructional materials to facilitate learning. However, both Mayer (2001) and Sweller (1999) acknowledge that there are some situations in which certain design principles should be applied with care. This position ultimately leads to design recommendations that have specific caveats. For instance, both recommend that designers offload textual materials to an audio format to improve learning, but caution that “there might be some circumstances where the presentation of textual materials in auditory form could be expected to have negative consequences” (Sweller, 1999, p.146). Both also recommend that designers present related information at the same time, but note that learning might also be improved when “students [are] able to listen to narrations without being distracted by related diagrams” (Mayer, 2001, p 104).

It is my contention that these design guidelines cannot be more definite due to a lack of precision in specifying the relationships between using multimedia instructional materials and cognitive processes important to learning. Extant theoretical approaches do not adequately specify how to-be-learned information is extracted and converted into a coherent mental model. Research suggests that people learn by forming mental codes to represent to-be-learned information and developing networks of meaningful associations among these mental codes to impart meaning to memories. Learning research suggests that it is important to design instructions in ways that foster verbal learning and the development of intellectual skills. Therefore, in this project, I explore the utility of a two stage process model for understanding how to configure multimedia instructional materials in ways that promote learning. To show that the present approach does not contradict the predictions made by other approaches to understanding how people learn from multimedia instructional materials, I also organize extant multimedia instructional design principles (Mayer, 2001; Sweller, 1999) according to whether they help learners perform each of the two stages in the process of learning from multimedia proposed here. Following this, I will explore a particular design guideline that seems to be the source of many special caveats, the modality principle, and show how the present approach can increase predictive power and thereby simplify related design guidelines.

1.4.1 Stage One: Creating mental codes to represent to-be-learned information

One approach to using an understanding of human cognitive architecture to examine how the configuration of multimedia instructional materials can influence learning is

Sweller's Cognitive Load Theory (CLT). The main tenant of this theory is that learning can be most effectively fostered if designers reduce extraneous cognitive load - working memory load attributed to the design of the instructional materials and not beneficial to learning (Chandler & Sweller, 1991) - in favor of promoting germane cognitive load - cognitive processes that are beneficial to remembering to-be-learned information (Sweller, van Merrienboer, & Paas, 1998). According to CLT, learners' ability to hold mental codes in working memory might be subject to limited capacity (c.f., Miller, 1956); well designed multimedia instructional materials cater to this limited capacity by configuring the materials in ways that reduce working memory demands, consequently freeing these resources to hold to-be-learned information in working memory (Mayer, 2001; Sweller, 1999).

Guidelines for the configuration of multimedia instructional materials are based on the limited capacity of working memory. For instance, giving learners a basic overview of the subject before a lesson (Mayer, Mathias, & Wetzell, 2002) might help learners devote cognitive resources to identifying and encoding only the ideas, concepts, and basic information that are important to the lesson (i.e., relevant to the subject). In addition, CLT predicts that fostering the consolidation of information into chunks and reducing the likelihood that learners must hold information from one medium in working memory as they search among other media (i.e., split attention) might help learners store information in working memory more efficiently, thereby reducing extraneous cognitive load (Chandler & Sweller, 1992; Sweller, 1999). As a consequence, learners are able to form cognitive structures (i.e., mental codes) to represent to-be-learned information. It is likely that facilitating this stage of learning is also reflected in improvements related to reducing information that might needlessly occupy working memory resources and thereby hinder encoding of relevant information, such as unnecessary words (Mayer et al., 1996), unrelated words and pictures (Harp & Mayer, 1998; Mayer et al., 2001), irrelevant sounds and music (Moreno & Mayer, 2000), and redundant materials (Kalyuga et al., 1999; Mayer et al., 2001).

In summary, CLT suggests that catering to learners' limited working memory capacity by reducing extraneous information and assisting learners as they chunk information improves learning. This is very similar to facilitating the formation of mental codes to represent to-be-learned information. According to the two stage process of learning from multimedia, these mental codes are the basic building blocks of learning. Configuring multimedia instructional materials in ways that help direct learners to view relevant parts of materials might be equivalent to facilitating the learners' ability to encode concepts, primarily leading to improved verbal learning outcomes.

1.4.2 Stage Two: Forming a network of associations among mental codes.

Extending the idea of improving learning by enabling learners to chunk to-be-learned information, Mayer's Cognitive Theory of Multimedia Learning (CTML) (2001) also uses an understanding of human cognitive architecture to examine how the configuration of multimedia instructional materials can influence learning. The main tenant of this *active processing* assumption is that meaningful learning requires the selection, organization, and integration of relevant to-be-learned information into a coherent mental model (Mayer, 2001). Following this logic, well designed multimedia instructional materials might help a learner recognize how concepts important to the lesson are associated with other concepts (Bishop & Cates, 2001), promoting the development of a comprehensive network of associations among facts.

Current guidelines for the configuration of multimedia instructional materials are devoted to making it easier for learners to identify how facts and concepts in different parts of the lesson (e.g., media) are related. In general, these guidelines suggest that learning can be improved by configuring multimedia instructional materials in ways that facilitate the integration of corresponding parts of different media. For instance, physically presenting corresponding text and pictures close to each other improves retention (Mayer, 1989; Moreno & Mayer, 1999), and transfer (Mayer, Steinhoff, Bower, & Mars, 1995; Sweller, Chandler, Tierney, & Cooper, 1990). Presenting corresponding words and pictures simultaneously rather than successively can foster retention (Mayer, Moreno, Boire, & Vagge, 1999), and transfer (Mayer & Anderson, 1991; Mayer et al., 1999; Mayer & Sims, 1994). Signals such as coloring, arrows, and icons help learners understand how to process complementary materials and improve problem solving transfer (Mautone & Mayer, 2001). Other methods of linking verbal segments to corresponding parts of the visual material, such as using an electronic flash (Jeung, Chandler, & Sweller, 1997), color coding (Kalyuga et al., 1999), or highlighting (Tabbers, 2002), have also been shown to improve learning.

In summary, CTML suggests that facilitating active processing leads to the formation of a complete mental model. Active processing can be described as forming associations between mental codes representing to-be-learned information. Configuring multimedia instructional materials in ways that help direct learners to identify how separate media correspond might be equivalent to facilitating the learners' ability to form associations between mental codes representing to-be-learned information, primarily leading to improved development of intellectual skills.

1.4.3 Using a second presentation modality

One last key assumption about human cognitive architecture made by CTML and echoed by CLT, is that learning can be enhanced when the configuration of multimedia instructional materials exploits humans' dual mode processing channels (c.f., Baddeley & Hitch, 1994; Paivio, 1986). A relevant empirical finding is that students learn better when visual text accompanying visual diagrams or animations is offloaded into audio narrations (Mayer & Moreno, 1998; Moreno & Mayer, 1999; Mousavi et al., 1995). Both CLT and CTML suggest that multimedia instructional materials used to present information in two modal forms (i.e., combining visual and auditory materials) activate segregated visual and auditory working memory subsystems that might not both be activated by unimodal materials (e.g., purely visual materials might only activate visual memory) (Ginns, 2005; Leahy, Chandler, & Sweller, 2003; Mayer, 2001; R. E. Mayer, 2005; Sweller, 1999, 2005). As a result of activating both auditory and visual working memories, multimodal presentation (e.g., combining pictures with narrations) increases functional cognitive resources compared to unimodal presentation (e.g., combining pictures with text). CLT implies that the activation of both auditory and visual working memories aids the accumulation of useful information (Brunken, Plass, & Leutner, 2004) and facilitates cognitive processing necessary to build schemas to be transferred to long term memory (Low & Sweller, 2005). CTML suggests that the activation enables learners to hold and process complementary information simultaneously (Mayer, 2001) as it is integrated into a coherent mental model (Richard E. Mayer, 2005). This accumulation of useful information in active memory makes it easier to combine it in long term storage and thus, is seen as the cause of the modality effect.

However, a recent meta-analysis of studies comparing learning from unimodal and multimodal presentations suggests that the modality effect (use of narrations leads to better learning than use of text) occurs only when presentations are matched to the time it takes for the auditory narration to play (Ginns, 2005). With extra time to examine instructional materials or when learners can control the rate of information, the use of text might aid learning because text can be more easily reexamined compared to narrations (Wickens & Hollands, 2000). In addition, when verbal materials are lengthy, using text might lead to better learning because narrations have the potential to produce more working memory demand than text (Sweller, 1999). It is possible that an extra acoustic memory trace automatically produced in response to spoken text (Penney, 1989) might increase extraneous cognitive load. Consequently, including narrations as part of instructional materials might burden working memory rather than functionally expand working memory. These facts suggest that the expansion of working memory explanation for modality effects could be incomplete (Tabbers, 2002).

An alternative way to explain the modality effect (how the presentation modality of verbal information influences learning) is to apply the two stage process model of learning from multimedia. This model leads to the hypothesis that the modality effect is a result of human's ability to *sense* information in multiple modalities (e.g., visual and auditory) at the same time (Mayer & Sims, 1994; Schnotz et al., 2002). That is, being able to perceive more information at once allows learners to increase the efficiency with which they can form mental codes to represent to-be-learned information because they can form two codes at once when possible or form one code to represent redundant information contained in both media at the same time. Eye movement research on the integration of pictures and text has found that subjects tend to read the text first and then look at the picture (Hegarty et al., 1992; Rayner et al., 2001). In multimodal presentations (with narration) learners do not need to read text before viewing the diagram. Therefore, the ability to sense more information at once might facilitate the acquisition of mental codes from both media simultaneously without having to, for example, read the text before being able to examine the diagrams. This allows learners in multimodal presentation conditions to form mental codes representing the to-be-learned information more quickly than learners in unimodal presentation conditions. When materials are not presented at an overwhelming pace, this is beneficial to learning.

The ability to sense more information at once might also allow the learner to visually search diagrammatic materials "online" (i.e., while listening to audio-narrations) and more easily locate points of correspondence in order to identify associations among complementary information (Mayer, 2001). As people listen to a story or follow instructions, they quickly move their eyes to those elements in an array that are most closely related to the words currently heard (Cooper, 1974). Therefore, being able to sense the two media at one time may foster the development of additional associations between previously established mental codes by allowing learners to search visual diagrams for elements that correspond to the part of the verbal narrations currently heard. However, learners cannot review narrations; they have one opportunity to learn the information from the verbal materials and how it is related to diagrammatic materials. This might create a situation in which learners in multimodal presentation conditions perform the process of forming associations among the information contained in different media in a qualitatively different way than learners in unimodal presentation conditions. That is, learning might follow the order of verbal presentation for narrations (because learners view the pictures in that order), but be more dependent upon learner differences in text versions (because there is less structure in how they view each alternately).

The proposed two-stage process model can be used to make specific predictions about the changes in learning caused by slowing the presentation pace of instructions. The model might predict that slowing presentation pace would not have much of an effect on learning from instructions using narrations, while it might have a positive effect on learning from instructions using text (see Figure 4). In contrast, the expanding working memory model (in both CLT and CTML versions) does not make any predictions about differential effect of slowing presentation pace on learning from instructions using narrations or text. The goal of this project is to compare the success of each model in predicting changes in the modality effect when presentation pace is slowed. In addition, the proposed project will analyze eye movement behavior to explore how different configurations of multimedia instructional materials are examined by learners. This analysis will correlate learners' eye movements with how learners form mental codes and identify how the information in separate media correspond (i.e., the two proposed stages of learning) while studying multimedia instructional materials.

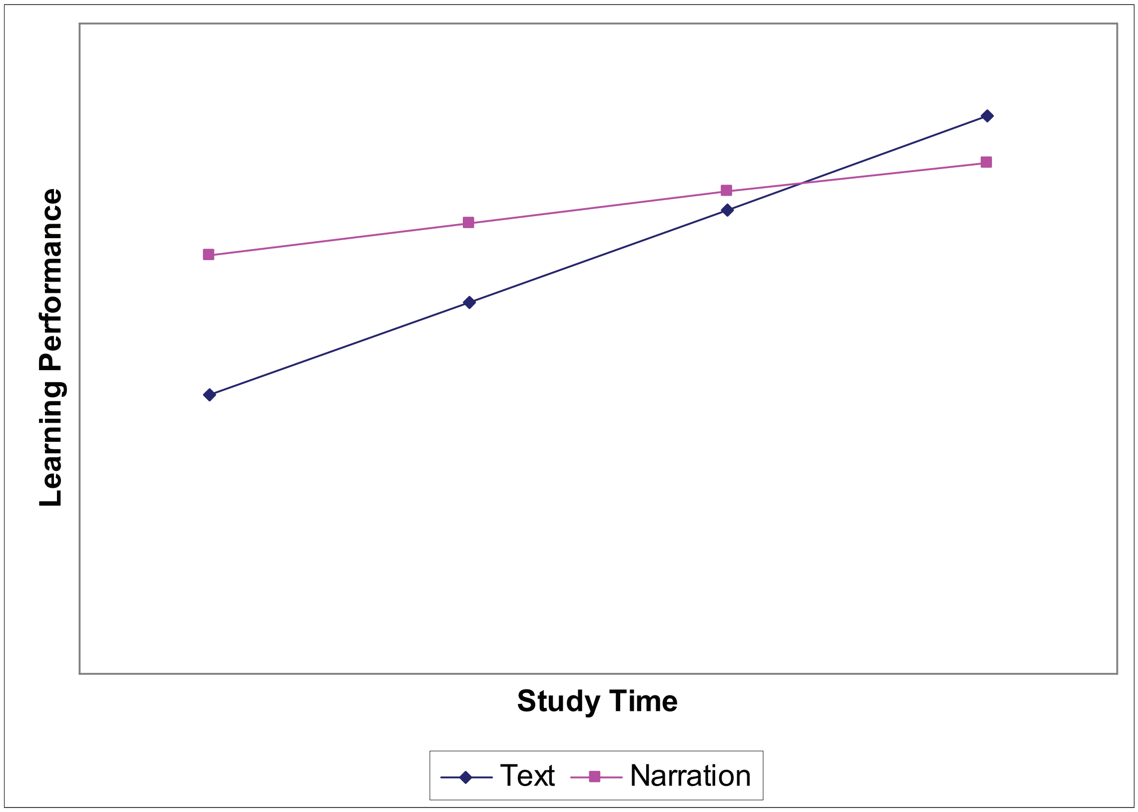


Figure 4: Schematic representation of predictions of proposed model when unimodal (text) and multimodal (narration) presentations are slowed.

If the two stage process model does a better job of predicting empirical results than the expanding working memory model, this understanding could lead to the development of more robust guidelines for when and why to mix modalities in multimedia instructional materials. Moreover, this approach to describing how people learn from multimedia instructional materials might extend to improving the understanding of other multimedia learning effects and improve other guidelines for how to configure multimedia instructional materials. The intended result is to allow designers to better understand how to configure multimedia instructional materials in ways that foster learning and how contextual factors such as available study time might influence this relationship.

1.5 Overview of the study

According to the two stage process model of learning from multimedia instructional materials, multimodal materials improve learning by exploiting the human's ability to sense information in multiple modalities (e.g., visual and auditory) (Mayer & Sims, 1994; Schnotz et al., 2002) and thereby help learners form cognitive structures to represent to-be-learned information contained in diagrams more quickly. Instead of replacing text with narrations, providing more study time might have the same effect (helping learners form cognitive structures to represent to-be-learned information by allowing more time to examine diagrammatic materials). The effects of being able to form cognitive structures to represent to-be-learned information should be most evident in increased dwell time on visual materials and verbal information learning outcomes. The proposed model also suggests that the use of multimodal presentation improves learning by enabling the learner to visually search diagrammatic materials while listening to audio-narrations and thereby more easily develop associations among to-be-learned information. Instead of replacing text with narrations, providing more study time might enable learners to systematically reexamine visual (both text and diagram) materials to identify how the information corresponds, and thereby develop more complex associations among cognitive structures. The effects of being able to develop complex associations among cognitive structures should be evident in the order of observing and the frequency of glancing at different parts of the visual materials as well as learners' performance on tests of intellectual skill development.

Comparisons of learning from presentations using different verbal presentation modalities are common in the literature (Ginns, 2005), but few also manipulate study time and even fewer systematically alter the presentation pace of self-paced materials (i.e., controlling how long each part of a lesson is presented) as a way to compare learning from various levels of study time. If both design configurations (mixing modalities and adding study time) improve learning (compared to unimodal presentations timed to match narration lengths) via similar mechanisms (i.e., facilitating the formation and association of mental codes representing the to-be-learned information), it might be predicted that as study time is increased, the benefits provided by the use of narrations instead of text (i.e., the modality effect) will be reduced.

In support of this suggestion, a recent study found the typical modality effect (transfer learning from narration > text) in conditions timed to match the length of narrations (19.3 minutes), but no differences in learning (transfer learning from narration = text) when study time was doubled (38.6 minutes) (Tabbers, 2002). This study manipulated both the modality (text versus narration) of verbal information and time on task for a web based multimedia lesson. A more recent study examined differences between groups of learners who, when able to pace information themselves, took more or less time to study (Harskamp, Mayer, & Suhre, 2007). The results were similar to Tabbers': there was the typical modality effect among learners who took less time to study (learning from multimodal instructions was better than learning from unimodal instructions), and a reverse modality effect on recall among learners who took more time to study (learning from unimodal instructions was better than learning from multimodal instructions). The results of both studies suggest that study time interacts with the effects of mixing modalities. Tabbers explained these results by suggesting that when "the pacing of instructions is such that learners have enough time to process them, visual text is at least as effective as spoken text" (2002, p.58). Though Tabbers did not specify how adding time to process text improves learning, it might be that the processing Tabbers has in mind is similar to allowing learners to direct their gaze to visual materials presenting information, enabling the formation of mental codes to represent to-be-learned information and/or allowing learners to reexamine these materials to facilitate the formation of associations among these mental codes.

In Tabbers' (2002) experiment, significant differences in learning were observed on transfer but not retention. This might be because the lesson topic was the development of a complex cognitive skill and therefore better design was more influential on cognitive skill (transfer) development. To adequately test outcomes associated with both verbal learning and the development of intellectual skills, my proposed research will utilize instructional materials that teach scientific cause-and-effect processes. Learning the topic included in these materials requires understanding both system states (i.e., verbal information) and relationships between those states (i.e., intellectual skill) (Mayer & Chandler, 2001). In addition, the instructional materials used in the present experiments contain more verbal materials than previous studies (i.e., easily identifiable verbal information).

To test verbal learning and the development of intellectual skills, the present study will use performance measures that independently tap verbal learning outcomes and intellectual skill learning outcomes. These measures are similar to dependent variables typically used in learning research (e.g., recall, transfer, matching, identification), but have been directly mapped to outcomes associated with two separable processes of learning from multimedia. The two-stage process model predicts that slowing presentation pace will have a positive effect on learning from instructions using text and much less of an effect on learning from instructions using narrations. A follow up study will compare conditions that lead to exceptional performance in each learning outcome using eye tracking methodology to match eye movement behavior related to the process of learning from multimedia instructional materials. The two-stage process model predicts that conditions that lead to better learning will also encourage eye movements that can be related to each of the proposed processes of learning from multimedia.

CHAPTER 2: EXPERIMENT 1

In Tabbers' (2002) study, when learning materials were presented in a self-paced manner (i.e., the learner could control the rate of presentation), there was a reverse modality effect on learning: performance on transfer questions following presentations with text was better than performance on transfer questions following presentations with narration. Those results bring into question whether there is a simple interaction between study time and the modality effect because learners in the self-paced condition (that led to a reverse modality effect) took about 24 minutes to study the materials (regardless of presentation modality). This was more time than provided when presentation was matched to the length of narrations (19.3 minutes; led to the typical modality effect), but less than the time provided in the longer condition (38.6 minutes; led to no modality effect). Though not examined statistically by the author, examining the group means[1] further reveals that changes in the provision of study time (i.e., adding time to study system-paced materials or allowing learners to take as much time as they liked with self-paced materials) did not increase mean learning performance following multimodal instructional materials (using narrations). However, changes in the provision of study time did increase mean learning performance following unimodal instructional materials (using text). On one hand, this might suggest that providing some extra study time can be a benefit to learning from unimodal instructional materials (using text), but excessive extra study time might be counterproductive. On the other hand, these results might also suggest that study time is not the only thing affecting performance; instead, providing

[1]Mean performance: audio system-paced, 63%; audio self-paced, 63%; visual system-paced: 51%; visual self-paced, 64%.

learners control over the pace of unimodal instructional materials (using text) increased learning performance. Moreover, it is possible that learners have more control over how they perform the two stages of learning from multimedia (e.g., deciding which materials to examine first or how often they switch attention to the other medium) when static materials are presented with ample study time.

This means that one potential reason for the pattern of results in Tabbers' study (2002) is the inherent difference in the way that learners were able to process visual-verbal (i.e., text) versus audio-verbal (i.e., narrations) presentations. During double-length multimodal conditions narrations were played twice consecutively; during self-paced multimodal conditions, each narration corresponding to the current instructional segment could be replayed in its entirety (from the start of the segment). Therefore, in Tabbers' multimodal presentation conditions, learners had less control over how they used the extra study time to examine verbal information compared to unimodal presentation conditions. That is, with text, learners could review the verbal information from the beginning *or* the middle *or* the end, but they were more constrained in how they could review narrations. To reduce the inherent differences in unimodal (i.e., text) versus multimodal (i.e., narrations) presentations, learners in the present experiment were able to rewind and fast forward narrations with a progress bar during all multimodal conditions (see Figure 5). Thus, like the text condition they could review the verbal information from any point. Providing control to the learner enabled learners to use available study time to reexamine the verbal information during narration conditions in a similar way to how they might reexamine the verbal information during text conditions. If utilized by learners, the ability to reexamine narrations has the potential to alter the effect of study time on learning from narrations accompanied by pictures (Zolna & Catrambone, 2007). However, if not used, additional study time might have little influence on learning from multimodal presentations.

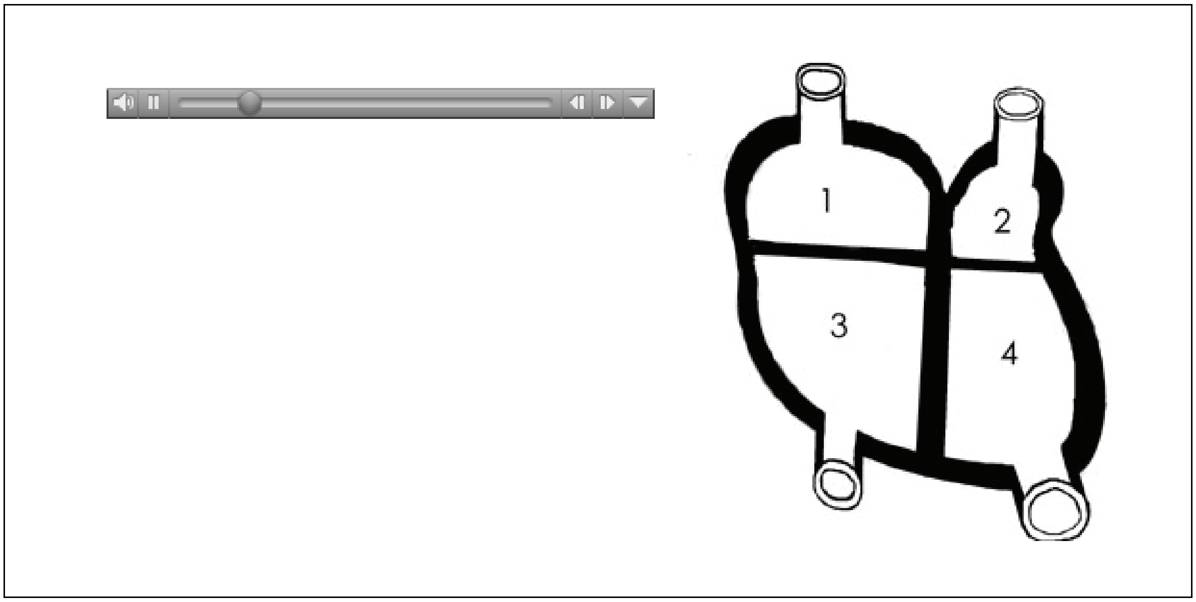


Figure 5: Narration condition of the heart lesson with progress bar in text area.

Harskamp et al. (2007) also compared performance of learners who took more or less time to study multimodal and unimodal instructional materials. In this study, the authors did not constrain presentation time, but provided learners the ability to replay presentations containing complementary words (text or narrations) and pictures as often as they wanted while measuring time taken to study the materials. This allowed them to perform a post-hoc grouping of participants into "fast learners" (who averaged 7.5 minutes of study time) and "slow learners" (who averaged about 11 minutes of study time). A modality effect was reported for the fast learners group. That is, transfer and overall test performance following multimodal presentation was better compared to transfer and overall test performance following unimodal presentation. For the slow learners there was a reverse modality effect. That is, recall test performance following multimodal presentation was actually *worse* compared to recall test performance following unimodal presentation. Those results suggest that doubling the presentation time is not necessary to alter the modality effect. Moreover, because learners who were in the slow condition were by definition studying materials longer, these results suggest that even if learners do utilize extra time to review materials, learning from multimodal presentation (pictures accompanied by narrations) might not improve.

In addition to reducing inherent differences in the ways that learners interact with verbal materials of different modalities, Experiment 1 used a more deliberate manipulation of study time than has been used in previous research. Harskamp et al. (2007) split learners into two post-hoc groups based on whether they took more or less time to study. Tabbers (2002) compared two system paced conditions, one matched to the length of narrations and one double the length of narrations, with a third self-paced condition. The manipulation of presentation pace used in the present experiment included four system-paced presentations. That is, pacing was always automatic, and never controlled by the learner. The levels of presentation pace were intended to examine learning from presentations that were paced to match the length of narrations, add a small amount of time, add an intermediate amount of time, and to double the provided study time. Comparing these conditions enabled me to more precisely examine the effect of study time on learning from multimodal and unimodal presentations.

The present materials have not been used in a comparison using visual and auditory presentation of verbal materials. Therefore, Experiment 1 also examined the generalizability of the modality effect and the influence of presentation pace to a new multimedia lesson: a lesson teaching the functioning of the human heart that lasts (at minimum) about 20 minutes and includes verbal materials accompanying 32 static visual diagrams. Furthermore, these materials have been shown to lead to robust learning due to integrating information from complementary verbal and diagrammatic materials (Butcher, 2006; Wolfe et al., 1998). Tabbers (2002) used a multimedia lesson illustrating worked-out examples to explain a procedure to design training for complex cognitive skills that lasted just under 20 minutes (under system paced conditions) and included verbal materials accompanying eight diagrams. Harskamp et al. (2007) used text and illustrations explaining animal behavior that took about six minutes (under system paced conditions). Therefore, the materials used here were longer and contained more total information, extending previous research to lessons of greater complexity.

Finally, the dependent measures used in the present study were specifically designed to test verbal information learning and intellectual skill development. Tabbers (2002) measured retention and transfer, but found an effect on transfer only. Harskamp et al. (2007) found differential patterns based on whether the test was recall or transfer. Each of the two learning outcomes used in the present study can be theoretically mapped to the two different processes of learning from multimedia. That is, performance in measures of recall can be theoretically linked to the formation of mental codes and verbal information learning; performance in measures of transfer can be theoretically linked to the development of associations between mental codes and intellectual skill development. In this way we could examine the effect of study time on learning from multimodal and unimodal presentations at two separate stages of learning.

With these materials, all theories to be compared in the present study (CLT, CTML and the two stage process model of learning from multimedia) predict that there will be a modality effect when presentation pace equals the time it takes to play the narration. This would also be similar to effects observed in Tabbers' (2002) system-paced condition, Harskamp et al.'s (2007) fast learners, and a pattern of effects commonly found in other comparisons of presentations manipulating the modality of verbal presentation (see Ginns, 2005 for review). However, each of the theories makes different predictions about how the changes in presentation pace influences learning from presentations using narrations and text. Therefore, each theory makes different predictions regarding the influence of presentation pace on the modality effect.

The two stage process model of learning from multimedia predicts that the modality effect would be reduced, eliminated, or even reversed when presentation pace is slowed. This is due to the relative ease with which text can be reexamined compared to narrations (Wickens & Hollands, 2000). Learners in unimodal conditions might steadily improve performance with slower presentation pace (which is equivalent to providing more study time) because it will make it easier and easier to examine and reexamine both diagrams and text. In contrast, learners in multimodal conditions might not be able to reexamine narrations as easily as they would text (as was the case in both Tabbers' self-paced and double-length conditions). Without being able to reexamine narrations, slowing presentation pace might have no influence on learning (i.e., forming mental codes or determining how information is related the). That is, learners can reexamine to-be-learned information from the text to form mental codes and/or form more associations among mental codes, but this might be too difficult with narrations - even when the ability to control them is provided. It might be possible that no amount of study time will make reexamining narrations easy enough for participants to actually benefit from slower paced materials.

The expansion of working memory hypothesis advanced by both CLT and CTML would predict that slowing presentation pace would not have different effects on learning from unimodal and multimodal presentations. According to the CLT view of the expansion of working memory, multimodal presentations increase total working memory capacity available to hold information from one instructional medium in working memory while searching among and integrating it with other instructional media (Mousavi et al., 1995). Slower presentation pace should allow learners more time to search among and integrate multiple physically distinct information streams more deliberately (i.e., in smaller sections), even if the learner has to use one sensory mechanism (such as vision for text and diagrams). That is, they would need to hold smaller amounts of information in working memory at any one time, increasing capacity dedicated to cognitive operations important to learning. According to CTML, the expansion of working memory promotes active processing by allowing learners to hold words and pictures in separate working memory stores at the same time (Mayer, 2001). Slower presentation pace should allow learners more time to select, organize, and integrate related materials and improve learning regardless of the modality of verbal materials (i.e., text or narrations). Therefore, learning would improve monotonically with a reduction in presentation pace whether instructional materials are unimodal or multimodal. Moreover, the expansion of working memory hypothesis does not distinguish between effects on verbal learning outcomes (probably related to selection and organization of information) or intellectual skill development (probably related to integration of materials).

2.1 Method

2.1.1 Participants

Two hundred and twenty four participants who received credit for participation in a Psychology class were distributed among eight conditions. An analysis of previous research in the domain of multimedia learning was conducted to determine group sizes.[2] There were 127 male participants and 97 female participants. Their age ranged from 16 to 26 years old (M = 19.43); 77 were freshman, 65 were sophomores, 40 were juniors and 39 were seniors. 88% were native English speakers and the remainder all reported that they spoke English fluently.

2.1.2 Materials

Heart lesson. Text and static diagrams describing the functioning of the heart were presented via 43 HTML pages presented in Mozilla Firefox. Web pages automatically advanced at a rate determined by condition (discussed in procedures section). The verbal materials were the simplest text about the heart and circulatory system used by Wolfe et al. (1998); this text was written at an elementary level and consists of 1,616 words. Each page contained between 1–4 sentences of verbal materials; 32 pages include simplified diagrams that depict concepts from the text. In unimodal

[2] When used in the past, the materials to be used in this study have been shown to be sensitive enough to show significant differences among 3 groups with 21, 22, and 24 participants on improvement in performance from pre- to post-test and in additional memory questions administered after the lesson, even with relatively small effect sizes ($\eta_p^2 = .10$). Using G Power, effect sizes from Butcher (2006) were computed to be approximately .33 and used to compute the necessary N to detect a significant difference in a global comparison and a special comparison with numerator df = 3 (to test effects of manipulating of study time). Other parameters used were an alpha = .05, power = .95 and groups = 8 (because Experiment 1 has eight between-subject manipulations). The analysis yielded a recommendation of 208 participants for the global comparison and 195 participants for the special comparison.

conditions, diagrams were presented adjacent to the text (see Figure 6). The multimodal version included concurrent narration that is equivalent to the lesson text spoken at a slow rate by a female voice. The narration began when the page with a corresponding diagram loaded. A progress bar and slide (similar to what one would see on any web-based video) allowed users to control the presentation of the narration by using the mouse to pause, rewind, fast forward or slide the progress indicator through the timeline (see Figure 5).

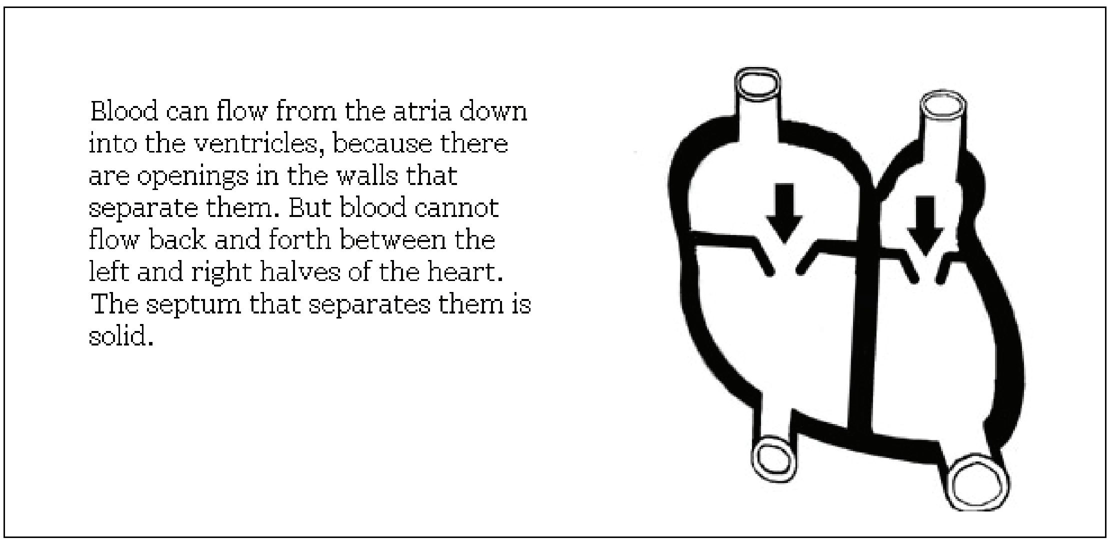


Figure 6: Example screen shot from the text condition of Experiment 1.

Individual differences measure. Prior to the experimental manipulation, the working memory capacity of participants was assessed using Automated O-Span. Automated O-Span (AOspan) is an automated version of a popular working memory capacity task (operation span) administered with E-Prime that takes about 15 minutes. This task has been shown to be a reliable and valid indicator of working memory capacity; dependent measures include an absolute span score shown to correlate with other measures of working memory (Unsworth, Heitz, Schrock, & Engle, 2005).

In this AOspan task, participants are asked to remember letter strings while performing simple arithmetic. The two tasks (arithmetic and letters to be remembered) are presented separately to participants, in an alternating manner. Participants indicate whether a given number is or is not the answer to an arithmetic problem throughout each trial and indicate recall of letter strings with the click of a mouse to indicate the letter and order they were presented at the end of the trial.

General learning test. The General Knowledge Test from Wolfe et al. (1998) was used to assess each participant's factual knowledge of general information about the human heart and circulatory system. The test consisted of 25 questions for 38 possible points. Ten total points are visually related, 15 total points are text related, 13 assess prior knowledge because they address information that is not included in the lesson . This test was administered both before and after the lesson to assess general learning.

Subjective measures. Participants completed the NASA TLX (Hart & Staveland, 1988) regarding the lesson (See Appendix C). The TLX measure self-reports of constructs such as cognitive load, physical load, and frustration. Summed together, the 5 scores yield an overall workload score. This subjective rating of workload may be informative in situations where increased cognitive load affects participants' ability to learn (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Participants also answered Likert-scale questions regarding the speed of the lesson, how they felt they performed on the post test, and their ability to identify how the materials correspond (see Appendix D).

Recall tests. Memory questions addressing specific details from the lesson about the heart and circulatory system were administered only after the learning phase. These questions assessed knowledge of facts and their basic relationships and were therefore appropriate to assess verbal learning outcomes (Gagne, 1972; Gagne et al., 2005). Questions were developed to include details that most participants would be unlikely to know or to be able to guess without having viewed the lesson. For example, "How many times will the average heart beat during a lifetime?" Some questions (e.g., "Where are the valves in the heart located?") asked participants to provide multiple pieces of information in their answers. Thus, a total of nine text-related memory questions (see Appendix A) were answerable for a maximum score of 17 points.

Picture-specific memory questions were also administered only after learning. These questions each included a diagram or part of a diagram from the lesson and asked participants to label parts of the diagram, fill in missing parts of the diagram, or to indicate what the diagram illustrates. Four picture-specific memory questions (see Appendix A) were answerable for a maximum score of 12 points.

Transfer tests. Inference questions related to the lesson were administered only after the learning phase. These questions assessed understanding of rules and principles inherent to the scientific system and relationships among parts of the system and were therefore appropriate to assess intellectual skill development (Gagne, 1972; Gagne et al., 2005). These tests required the participant to integrate information found in the lesson and to apply such information to new situations or problems; answers were not addressed explicitly in the learning materials. For example one question was, "What would be the consequences of a large hole in the septum that separates the left and right ventricles?" Correctly answering this question required the participant to recognize that the septum exists to separate the oxygenated and deoxygenated blood (a fact not explicitly addressed in the lesson) and to apply this structural information to the learned concepts of energy production and carbon dioxide removal. Full integration of the learning materials would allow the participant to reason about how a hole in the septum would affect essential processes (e.g., less oxygen in the blood means less energy, too much carbon dioxide can kill body cells) and to explain the resulting effects on the human body (fatigue, possible death, etc.). Some questions were picture-specific, some text-specific. The five picture-specific inference questions were answerable for a maximum score of 11 points and five text-specific transfer questions were answerable for a maximum score of 8 points (see Appendix A).

2.1.3 Procedure

The study design was a 2 (modality of verbal information) x 4 (study time) between subjects factorial design with each participant tested immediately following the lesson (immediate test). Additionally, they were tested after a 7 to 9 day retention interval (delayed test). Modality of verbal information was manipulated to be either text or narration. Study time was system paced for all conditions. However, the pace of presentation was manipulated under four conditions: timed to narration length (standard), presented for an extra 3 seconds narration length per page (standard plus 3), to be displayed for 50% longer than the standard paced condition (150%), to be displayed for 100% longer than the system paced condition (200%). Participants were informed that the eight people who did best on the post lesson test would be awarded $10.

Dependent measures accounted for general learning improvement (from pre-test to post-test, based only on questions in the general learning test that were addressed in the lesson), recall (summation of text- and picture-specific recall questions), and transfer (summation of text- and picture-specific transfer questions).

Upon arriving to the laboratory, participants read and signed an experimental consent form. They were then briefed in a group (up to nine participants at a time) that they are to view a lesson on the functioning of the human heart and answer some questions about the lesson. Participants were seated at their own desk with headphones and a computer. They were given a paper and pencil background questionnaire (See Appendix E), pre-test, and the Automated O-Span and then allowed to begin the lesson on their own. The lesson advanced according to the pacing condition; learners in the narration condition were able to use the mouse to control the narration. After observing the lesson they completed a paper and pencil NASA TLX, general learning post-test, recall test, transfer test, and the subjective Likert questions.

After the participant completed the experiment, a return appointment was scheduled for the delayed tests. When the participant returned, they were seated at a desk with instructions to complete the tests and allowed to leave when they had done so.

Coding data. Data were scored by three coders who were blind to treatment condition. Two coders rated each data point to check for consistency as necessary. To determine the necessity for a second rater, one primary coder rated about one third of the total data (data from 75 participants) and two secondary coders rated half those data (about 40 each). An individual question that was shown to have high inter-rater reliability (an individual question that had over 90% agreement: 7 or less discrepancies among the 75 sets of data) was deemed to no longer need to be rated by two separate coders. The three coders then met to clarify and discuss the questions with less than 90% accuracy. Following this meeting, the primary coder again rated about one third of the total data (data from 70 participants) and two secondary coders rated half those data (about 35 each). Again, an individual question that was shown to have high inter-rater reliability (an individual question that had over 90% agreement: 7 or less discrepancies among the 70 sets of data) was deemed to no longer need to be rated by two separate coders. The three coders met to clarify and discuss questions with less than 90% accuracy. Following this meeting, the primary coder again rated the remainder of the data (data from 79 participants) and two secondary coders rated half those data (about 40 each). The three coders met again to discuss discrepancies.

2.2 Results and Discussion

To assess the efficacy of the data gathered, the heart test was assessed for internal consistency reliability using Cronbach's alpha. This measures how well a set of items measures a single unidimensional latent construct (i.e., learning from the previous lesson). The heart test had a reliability of 0.80. A reliability of 0.70 is commonly regarded as an acceptable reliability coefficient (Nunnaly, 1978), so no items were deleted. The average Total OSpan score of participants was 62.12 (SD = 10.495), average GPA was 3.16 (SD = .53). The average number of biology courses taken by participants was 1.66 (SD = 1.23), 143 last took biology in high school, 67 in college; 27 participants reported that they were pre-med.

Before analysis, outliers were deleted. The removed outliers were those participants whose total raw score (summed scores) for the heart lesson was more than 2 standard deviations from the mean. Fifteen subjects were removed from analysis by this method, no more than three in any one condition. This left between 25 and 29 participants in each condition. No remaining subjects had a total raw score on the retention test that was more than 2 standard deviations from the mean.

2.2.1 Learning

Manipulation Check. The lesson, in all configurations, was successful in teaching learners about the functioning of the heart and circulatory system: dependent sample t-tests showed reliable improvement in performance on the general knowledge test from pre- to post-test for all conditions (see Table 2). Correlations between scores were significant and strong, indicating that those who did well on the pre-test also did well on the post-test.

Table 2: Mean performance for Pre- to Post-learning phase and t-test of differences for heart general knowledge test performance as a function of verbal presentation modality and presentation pace.

Modality	**Presentation Pace**	***N***	**Pre-Test (*SD*)**	**Post-Test (*SD*)**	**Correlation**	**t-value**
Text						
	Standard	26	10.85 (6.23)	20.23 (5.05)	.593*	9.215*
	Standard+3	25	11.60 (7.05)	21.16 (3.65)	.638*	8.702*
	150%	26	10.50 (6.65)	21.04 (4.46)	.566*	9.730*
	200%	24	13.58 (8.93)	23.54 (5.94)	.805*	8.966*
Narration						
	Standard	26	13.12 (5.84)	22.69 (3.46)	.673*	11.234*
	Standard+3	29	12.38 (7.91)	22.41 (4.27)	.749*	9. 378*
	150%	25	12.10 (7.83)	21.56 (3.73)	.721*	8.224*
	200%	25	10.48 (9.66)	21.48 (5.19)	.875*	9.648*

Note: SD = standard deviation. Maximum possible score = 38. SD = standard deviation. * = p < .05. Condition types described in methods.

Overall learning. To assess the experimental hypotheses (i.e., assess the influences of verbal presentation modality and presentation pace on immediate and delayed learning) between subjects univariate ANOVA were conducted with modality of verbal presentation (audio vs. narration) and presentation pace (four levels) as fixed factors. Separate univariate ANOVAS were conducted to avoid the pitfalls of using MANOVA with multiple dependent variables that might not be independently sampled. Four separate analyses, one with each test performance measure (immediate recall, immediate transfer, delayed recall, and delayed transfer) as dependent factors were performed.

The analysis of performance on immediate recall revealed a marginal statistically significant main effect of presentation pace with a medium effect size on immediate recall performance, $F(3,186) = 2.36$, $p = .07$, $\eta_p^2 = .04$, no main effect for verbal presentation modality, $F(1, 186) < 1$, and no interaction between the two, $F(3,186) = 1.70$, $p = .17$. See Figure 7. The analysis of performance on immediate transfer revealed no significant main effects of presentation pace, $F(3, 186) < 1$, , no main effect for verbal presentation modality, $F(1, 186) < 1$, and no interaction between the two, $F(3, 186) < 1$. The analysis of performance on delayed recall revealed no significant main effects of presentation pace, $F(3, 186) < 1$, no main effect for verbal presentation modality, $F(1, 186) < 1$, and no interaction between the two, $F(3, 186) < 1$. The analysis of performance on delayed transfer revealed no significant main effects of presentation pace, $F(3, 186) = 1.54$, $p =.21$, no main effect for verbal presentation modality, $F(1, 186) = 1.73$, $p =.19$, and no interaction between the two, $F(3, 186) < 1$. To explore which group mean differences (see Table 3 for group means) were contributing to the main effect of presentation pace on immediate recall performance, post-hoc LSD comparisons between levels of presentation pace and collapsed over presentation modality were examined. These comparisons showed that participants who received the standard+3 pace performed significantly worse on the recall questions immediately after the lesson (*M* = 17.93, *SE* = .582) than participants in the group who received 150% pace (20.01, .583), *p* = .01 or 200% pace (19.78, .585), *p* = .03. No other reliable pairwise differences were observed.

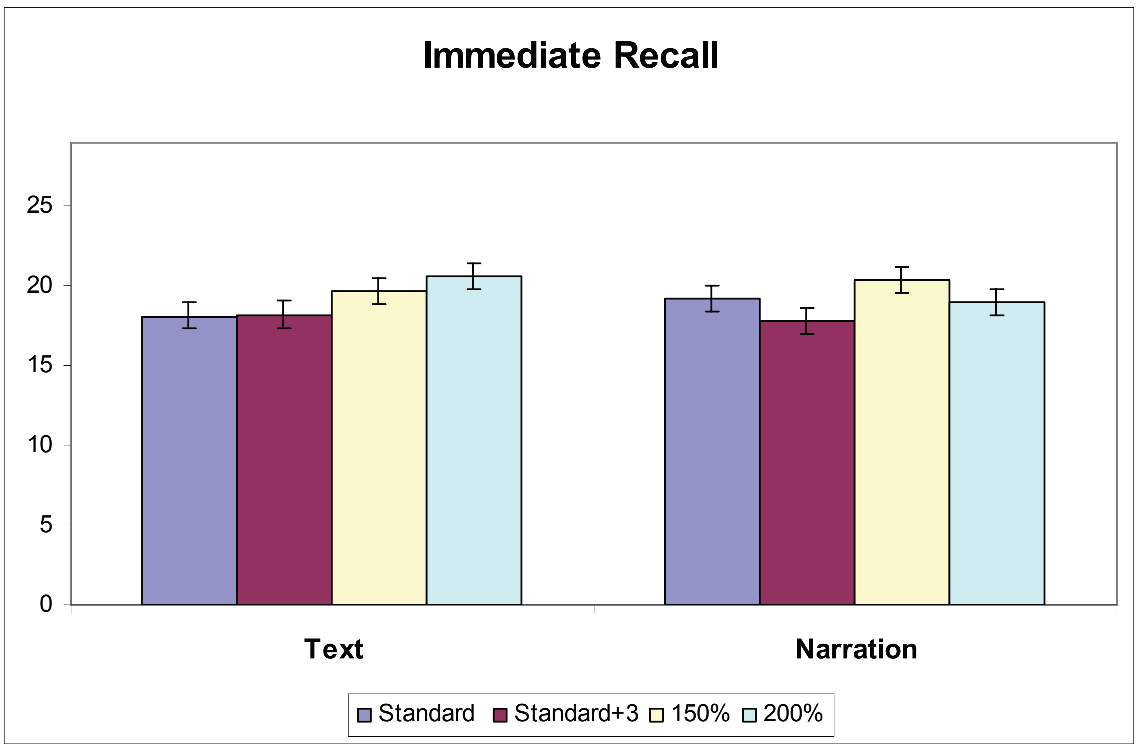


Figure 7: Chart depicting immediate recall performance following the heart lesson.

Table 3: Mean performance on each test administered immediately (immediate test) and 7 to 9 days after (delayed test) after the heart lesson as a function of verbal presentation modality and presentation pace.

			Immediate Test		Delayed Test	
Modality	**Presentation Pace**	***N***	**Recall (*SD*)**	**Transfer (*SD*)**	**Recall (*SD*)**	**Transfer (*SD*)**
Text						
	Standard	25	17.66 (5.14)	8.40 (3.38)	16.96 (4.86)	8.61 (2.28)
	Standard+3	23	18.20 (3.95)	8.63 (2.26)	17.78 (3.81)	8.22 (1.87)
	150%	25	18.96 (3.77)	7.88 (1.84)	17.20 (4.03)	7.76 (2.61)
	200%	26	20.63 (4.61)	8.46 (2.36)	16.69 (5.45)	7.98 (3,08)
Narration						
	Standard	26	19.21 (3.41)	8.96 (2.51)	16.31 (4.72)	8.58 (3.14)
	Standard+3	29	17.78 (5.06)	8.16 (2.56)	16.95 (5.60)	7.28 (2.62)
	150%	24	20.40 (3.35)	8.19 (2.77)	18.71 (3.84)	7.42 (2.46)
	200%	23	18.83 (4.07)	7.72 (2.88)	17.57 (4.43)	7.11 (3.46)

Note: Maximum recall score = 29; Maximum transfer score = 20. SD = standard deviation.

The results of these ANOVAs suggest that the only marginally reliable differences detected in the present experiment were in performance on immediate recall questions, and no reliable differences in performance on transfer questions were detected. This might suggest that there was no measure used in this experiment that was adequate for detecting differences in deeper learning or there were no such effects based on manipulating presentation format or pace. I suggest that this lack of reliable effects related to transfer performance was due to the fact that the materials used were focused on recall information (e.g., facts and terms). These materials were not designed to produce detectable differences in the formation of associations among those facts (i.e., developing robust mental models). Past experiments including these materials have also failed to detect differences in performance on transfer questions (Butcher, 2006); researchers cited difficulty in making necessary inferences and suggesting that questions requiring fewer inferences with more direct transfer opportunities would have resulted in better comprehension data. After addressing this issue by adding more transfer questions to the tests used in the cited experiment and using at least one manipulation (verbal presentation modality) shown to have reliable and detectable effects on transfer test performance (Mayer & Moreno, 1998; Tabbers, 2002), differences in performance on transfer questions were still not detected. This suggests that the lack of an effect on transfer might be due to the lesson content and not the measures of learning or manipulations.

In addition to this lack of an effect on immediate transfer questions, the present experiment did not detect reliable differences on delayed recall even after detecting an effect on immediate recall. Delayed recall measures often detect similar aspects of learning that are measured by immediate transfer questions (Schmidt & Bjork, 1992). This result further supports the supposition that this lack of reliable differences in transfer test performance and delayed recall test performance could be the result of the materials rather than the test or manipulations. In addition, Levene's test revealed that the error variance of immediate transfer questions was not equal across groups, $p = .05$, suggesting that participants' performance on the immediate transfer test was systematically different than performance on the recall test (which had equal variance across groups). Therefore, I conclude that the failure to detect differences in transfer performance might be attributed to the validity of the test rather than the experimental conditions. That is, the test did not do a good job of eliciting quality responses from participants, and therefore did not adequately assess their learning. Due to this apparent artifact of the materials, the rest of the data analysis will focus only on results related to immediate recall.

Presentation pace. According to the two-stage model of learning from multimedia, presentation pace was expected to have an influence on learning. That is, more time to examine the materials should allow learners to perform stages important to learning, leading to a linear effect of presentation pace on learning. To further explore this *a priori* hypothesis, I performed a one-way ANOVA with presentation pace as the independent factor and immediate recall performance as the dependent factor. There was a significant effect of presentation pace on immediate recall performance, $F(3,205) = 2.64$, $p = .05$, $MSE = 17.86$. Moreover, there was a significant linear pattern, $F(1,205) = 5.67$, $p = .02$, $MSE = 17.86$, of improving performance as presentation pace was slowed (Table 4).

Table 4: Mean performance on the immediate recall test for each level of time collapsed across presentation modality.

Presentation Pace	***N***	**Immediate Recall (*SD*)**
Standard	52	18.33 (4.42)
Standard + 3	54	18.02 (4.51)
150%	51	19.78 (3.59)
200%	52	19.82 (4.30)

Note: Maximum score = 29. SD = standard deviation.

Based on past research and the proposed model, it was predicted that when verbal materials are presented via text, learning would improve as presentation pace is slowed. It was hypothesized that this is due to the additional time provided for learners to form and integrate mental codes representing to-be-learned information from simultaneously presented visual media (i.e., text and diagrams). However, it was predicted that this improvement might not be as strong when verbal materials are presented via narration because no extra time is needed to examine diagrams in multimodal presentations (diagrams can be attended during the entire presentations) or determine their relationships with diagrams (this is done immediately, while listening to narrations). To separately explore these hypothesized differential effects of presentation pace on learning from versions of the heart lesson using text and narration, two separate follow up one-way ANOVAs were performed, one including the group that received narrations and one including the group that received text. This analysis revealed a significant linear pattern of improving immediate recall performance as presentation pace was slowed among text conditions, $F(1, 99) = 7.27$, $p = .01$, $MSE = 19.35$, but no such pattern among narration conditions, $F(1, 102) < 1$. This suggests that the slowing of presentation pace facilitated immediate recall performance following lessons with text, but not lessons with narration, confirming the hypothesis that slower presentations improve learning from unimodal but not multimodal presentation. This could contribute to the fact that the modality effect is rarely observed when presentation pace is not matched to the time it takes to play narrations (Ginns, 2005) and is in line with the trends of previous studies reviewed in this paper that have compared the modality effect across conditions of varying presentation pace (Harskamp et al., 2007; Tabbers, 2002).

According to the framework suggested in this paper, slowing of presentation pace of presentations with text even a small amount should improve the formation of mental codes due to providing learners the ability to more completely examine the diagrammatic materials after reading the text. This should improve recall. Furthermore, the theory suggests that further slowing the pace of presentation should enable learners to take advantage of the static nature of text and reread verbal information as well as switch back and forth among media, determining their relationship and forming associations between mental codes. Therefore, it was predicted that an increase in recall performance would come with a slightly slower presentation pace and increases in transfer performance would come with much slower presentation paces. Further examination of mean differences within the unimodal conditions using post-hoc LSD comparisons shows that the only reliable pairwise difference was between the fastest (standard) and slowest (200%) presentation pace, mean difference = 2.77, $p = .02$. These results suggest that there were some benefits on recall of providing much slower presentation with text materials. Other research in the domain of learning suggests that information that involves a variety of associations can be better remembered (Anderson, 1995; Craik & Lockhart, 1972), and when learners integrate information from two or more media they are able to remember more information than from individual media (Bransford & Franks, 1971). It might be that the formation of associations between mental codes can also improve recall, and in the absence of a sensitive transfer measure this indicates that slower presentation pace can improve learning from multimedia.

The two stage process model of learning from multimedia also suggests that learners are able to easily form and integrate mental codes representing to-be-learned information from both media in conditions using narrations even with fast presentation pace (standard condition). Therefore, the present framework would predict that a smaller reduction in presentation pace (i.e., standard+3 condition or 150% condition) would not lead to reliable improvements on learning performance of any type. However, it was predicted that when presentation pace was slowed enough so that learners could reexamine narrations similarly to how they would reexamine text (e.g., 200% condition), there might be some learning improvements. The previously reported one way ANOVA suggests that there was no consistent relationship between increasing study time and learning from lessons using narrations. It is possible that learners were not able to take advantage of extra study time to reexamine narrations. If this is the case, that would explain the lack of an effect of slowing presentation pace among conditions with narrations and further work needs to be done to determine how narrations can be presented in ways that they can be reexamined in cases where there is opportunity to provide learners more study time.

Modality effect. Due to the differential effects of presentation pace on learning from presentations using text versus narration, it was expected that the modality effect would be reduced and possibly reversed as presentation was slowed. However, the MANOVA reported above indicated that there were no reliable differences in learning as an effect of verbal presentation modality. To further explore the modality effect and the possibility of its change with presentation pace, I examined performance on the general knowledge test. This test was almost exclusively recall questions and was therefore an appropriate way to explore learning outcomes of various presentations of the heart lesson. However, not all of the information on the test was included in the lesson, therefore, only questions addressed in the lesson were used in the following analysis (see questions marked 'relevant' in Appendix A).

To explore the modality effect, independent sample t-tests with presentation modality as a grouping factor and performance on the relevant questions included in the general knowledge test as a dependent factor were performed at each level of presentation pace. In line with the modality effect, among standard paced conditions participants in narrations conditions answered more questions correct (mean difference = 2.26) than those in text conditions, $t(52) = 2.40$, $p = .02$. There was no reliable difference among standard+3 conditions, $t(50) = 1.05$, $p = .30$, 150% conditions, $t(49) = .67$, or 200% conditions, $t(49) = .90$. However, there was evidence of a reversal of the modality effect among the 200% conditions, with a reversal of the rank order of mean performance following text versus narration (mean difference = 1.02, see **Table 5**).

Table 5: Mean score on the post-lesson general knowledge test questions that tested information contained in the lesson; mean difference and the critical t-value testing differences between verbal presentation modality at each level of presentation pace.

Presentation Pace	Text		Narration		Mean	t-
	N	M (SD)	N	M (SD)	Difference	value
Standard	26	19.23 (4.023)	26	21.50 (2.672)	2.269	2.369*
Standard+3	25	20.20 (3.304)	29	21.172 (3.44)	0.972	1.054
150%	26	20.35 (4.009)	25	20.84 (3.387)	0.494	0.672
200%	26	21.42 (4.002)	25	20.40 (3.032)	1.023	0.896

Note: Maximum score = 26. SD = standard deviation. * = $p < .05$.

These results are similar to those of Tabbers (2002) and Harskamp et al. (2007). The present results extend these past studies with a deliberate manipulation of presentation pace. This manipulation was intended to examine the effect of presentation pace on learning from instructions presenting verbal materials with text and narrations. In this way, the present experiment was used to explore the modality effect at different presentation rates, addressing the recent findings that the modality effect occurs only when presentations are paced to match the time it takes to play the narrations (Ginns, 2005). The findings described here support the suggestion that the modality effect is reduced when presentation is not matched to the length of narrations. This combined with the previous analysis suggest that this is partially due to the fact that reducing the pace of presentation has a benefit on learning from materials using text and not on learning from materials using narrations. These findings expand upon Tabbers' (2002) suggestion that a reversal in the modality effect is related to learners' ability to use extra study time to process unimodal materials.

2.2.2 Subjective experience

Subjective measures of the learning experience were also administered after the lessons. Table 6 shows mean responses on subjective ratings of subjective mental workload taken by the NASA-TLX. A univariate ANOVA examining the effect of the two presentation manipulations (modality of verbal presentation and presentation pace) on subjective mental workload showed a main effect of presentation pace, $F(3, 201) = 4.16$, $p < .01$, $\eta_p^2 = .06$, and no main effect of presentation mode, $F(1, 201) = 1.69$, $p = .20$. There was no interaction between the two manipulations, $F(3, 201) = 1.79$, $p = .15$. Overall subjective mental workload was higher for shorter presentation times with post-hoc LSD comparisons showing subjective mental workload significantly less with 200% presentations pace than standard+3, $p = .02$, or standard presentation pace, $p < .01$, but not significantly different based on presentation modality.

However, on Likert questions very few participants rated presentations as too fast to learn from (see Table 7). The majority, 83%, reported that the presentation rate was fine or they would make it faster. Less than half (10 out of 25) of those who got the fastest text presentations reported that they would have slowed the rate of presentation if they could, and about one-third (9 of 26) of those who got the fastest narration presentations would have slowed the rate of presentation if they could. Examining the means for both this Likert scale question and the NASA-TLX Temporal Demand subscale, it appears that subjective ratings of the presentation being too fast (i.e., wanting to slow the presentation or reporting high temporal demand) can be alleviated by small amounts of extra time for text presentations (adding three seconds per slide), but requires more extra time (adding 50%) in narration conditions. Interestingly, the slowest presentations seemed to elicit similar reactions (nearly equivalent mean ratings on both questions).

Table 6: Mean NASA-TLX scores after the heart lesson for each subscale and the summed total as a function of verbal presentation modality and presentation pace.

Modality	**Presentation Pace**	***N***	**Total (SD)**	**Mental Demand (*SD*)**	**Physical Demand (*SD*)**	**Temporal Demand (*SD*)**	**Perform-ance (*SD*)**	**Effort (*SD*)**	**Frustra-tion Level (*SD*)**
Text									
	Standard	25	221.4 (96.99)	55.20 (22.01)	7.40 (12.59)	44.00 (26.26)	42.20 (22.04)	54.60 (18.42)	39.80 (25.35)
	Standard+3	23	245.38 (79.96)	47.00 (23.67)	6.44 (13.94)	34.60 (24.28)	26.80 (13.30)	47.00 (20.57)	23.40 (21.83)
	150%	25	187.20 (87.26)	43.21 (29.00)	6.79 (7.96)	27.14 (23.90)	31.96 (17.02)	49.29 (25.41)	31.25 (22.96)
	200%	26	186.15 (97.35)	47.04 (22.33)	2.59 (4.01)	20.37 (18.00)	26.30 (25.40)	47.22 (22.89)	25.00 (21.79)
Narration									
	Standard	26	243.20 (86.34)	49.42 (22.55)	5.58 (8.29)	43.27 (27.82)	40.19 (20.17)	48.27 (22.71)	36.54 (26.34)
	Standard+3	29	185.24 (85.60)	54.64 (23.88)	8.57 (14.52)	46.79 (21.22)	37.68 (21.06)	56.79 (22.08)	41.61 (25.75)
	150%	24	195.00 (82.35)	40.38 (23.02)	4.81 (6.24)	31.35 (23.56)	30.96 (20.79)	50.00 (25.46)	29.04 (26.65)
	200%	23	166.92 (65.44)	48.85 (25.35)	8.46 (16.96)	19.62 (18.27)	31.92 (19.45)	39.62 (27.60)	37.69 (32.99)

Note: Maximum score for each subscale = 100; Maximum score for total = 600. SD = standard deviation.

Table 7: Mean responses on the heart lesson related Likert questions as a function of verbal presentation modality and presentation pace.

Modality	Presentation Pace	*N*	Speed (*SD*)	Understanding *M* (*SD*)	Related *M* (*SD*)	Performance (*SD*)
Text						
	Standard	25	3.28 (0.89)	2.20 (0.87)	1.88 (0.83)	2.76 (0.93)
	Standard+3	23	2.65 (0.67)	1.92 (0.56)	1.62 (0.75)	2.54 (0.91)
	150%	25	2.68 (0.86)	1.89 (0.63)	1.57 (0.74)	2.75 (0.80)
	200%	26	2.11 (0.51)	1.78 (0.70)	1.63 (0.74)	2.30 (0.82)
Narration						
	Standard	26	3.04 (0.82)	2.04 (0.77)	1.81 (0.69)	3.08 (0.94)
	Standard+3	29	3.11 (0.83)	2.04 (0.84)	1.68 (0.77)	2.61 (1.10)
	150%	24	2.65 (0.89)	1.88 (0.71)	1.81 (1.06)	2.38 (1.13)
	200%	23	2.12 (0.73)	2.04 (0.89)	1.60 (.577)	2.64 (1.08)

Note: Each score is out of 5. SD = standard deviation. See Appendix D for items.

This experiment was an attempt to explore the influence of verbal presentation modality and presentation pace more deeply through the exploration of theoretical stages of learning and their relation to different types of learning outcomes. Unfortunately the materials used might not be adequate to detect reliable differences in anything but tests of immediate recall. In the next experiment, I used a different set of learning materials intended to detect reliable differences in more that just immediate recall. Based on this more detailed data gathering, including types of learning, verbally versus visually related materials, and a more deliberate manipulation of study time, these results will help examine the two stage process model of learning from multimedia instructional materials.

Analysis of this model can help explain how learners process unimodal and multimodal instructional materials when they have more study time.

CHAPTER 3: EXPERIMENT 2

The goal of Experiment 2 was to examine the interaction between modality and presentation pace for animated materials (as opposed to the static materials used in Experiment 1). The lesson in Experiment 2 used a 16 segment animation (each segment includes a sentence or two and one corresponding visual event) to teach how lightning is formed in (at minimum) about one minute and 45 seconds. The experimental manipulations replicated Experiment 1 (manipulating verbal presentation modality and presentation pace) to test the predictions made by the two stage process model of learning from multimedia. That is, the theory was tested on a lesson with a shorter learning course, a different topic, and with different outcome measures. The materials used in this experiment included complementary dynamic animations and verbal materials that prior research has shown to be sensitive to manipulations of modality (Mayer & Moreno, 1998) and altering the pace of presentation (Mayer & Chandler, 2001). In addition, these materials have been demonstrated to detect reliable differences in performance when including transfer tests (Mayer & Moreno, 1998).

Manipulating animated instructional materials to increase study time could be done in at least two ways. One option is adding time between segments of the instructions (cf., Mayer & Chandler, 2001). Implications include narrations matching the animation the same way regardless of presentation time, but periods of inactivity between segments (for conditions with text and with narrations) that might not be used for learning. Another option is slowing the animation and beginning the narrations as its corresponding visual event begins. Implications include the narration finishing before the animation of each segment, but learners are more likely to use the entire presentation time to study the materials, regardless of condition. In this experiment the latter option of slowing animation was chosen to encourage learners to use the entire time allotted for studying the materials and avoid the loss of concentration that might accompany a long pause in materials.

Research is inconclusive as to how learning differs when using motion pictures or animations instead of static graphics in instructional materials (Byrne et al., 1999; Hegarty, 2004). It has been suggested that motion pictures place increased demands on learners because they are transient and previous states must be held in memory if they are to be integrated with new knowledge (Stenning, 1998). Slowing the presentation pace might not adequately reduce the amount of information the learner must hold in working memory at any one time, regardless of the presentation modality of verbal materials. However, providing learners more study time to select, organize, and integrate related materials would also benefit learning regardless of the modality of verbal materials (i.e., text or narrations) by allowing time for active processing. For these two reasons, the expansion of working memory hypothesis advanced by both Cognitive Load Theory and Cognitive Theory of Multimedia Learning would predict that slowing the presentation pace of animated instructional materials would not have different effects on learning from unimodal and multimodal presentations.

According to the two stage process model of learning from multimedia, the modality effect in presentations using animated visual materials can be attributed to the same source as modality effects in presentations using static visual materials: multimodal presentation allows parallel sensation of two distinct media and the ‘online’ integration of complementary information contained in those materials. However, increasing the time that animated multimedia materials are presented might influence these benefits differently compared to when visual materials are static. That is, the two stage process model of learning from multimedia suggests that the two main manipulations in this experiment (verbal presentation modality and presentation pace) might interact and have different outcomes for verbal information learning (e.g., recall test performance) and intellectual skill development (e.g., transfer test performance) when accompanying visual materials are animated. That is, the interaction of effects is predicted to be different than when instructions include static visual materials.

When instructions include animated materials, slowing presentation pace is likely to aid the first process of learning from multimedia, the formation of mental codes to represent to-be-learned information (as with instructions using static visual materials). It is predicted that slower pacing might be especially beneficial with animated instructions because this will reduce the negative effects of their inherent transience; learners will have ample time to form mental codes to represent the to-be-learned information and are less likely to miss the important part of a visual event. Therefore, recall learning should be improved when presentation of the lesson is slowed. However, when visual materials are animated it is likely to take a significant amount of slowing (compared to when visual materials are static) to improve learning. Because learners can listen to narrations while they view animations, it is likely that smaller reductions in pace will be necessary to see this improvement when verbal materials are presented via narrations than when presented via text.

However, slowing presentation pace in this way is predicted to have very different effects on determining the relationships among to-be-learned information, the second stage of learning from animated multimedia instructional materials. On one hand, slowing animation accompanied by narrations might actually reduce learners' ability to form associations among to-be-learned information because the alteration in pace might reduce the likelihood that corresponding verbal and visual materials are presented at the same time. Presenting corresponding words and pictures simultaneously (at the same time) rather than successively (at different times) can foster retention (Mayer et al., 1999). Therefore, decreasing the pace and altering the temporal contiguity of corresponding information, even when they are multimodal, might make it more difficult to form associations among corresponding information contained in the two distinct media. On the other hand, slowing animation accompanied by text might encourage the formation of associations between mental codes by encouraging learners to switch back and forth between the animation and text to learn all the information contained in the two distinct media. This might force learners to search for points of correspondence and identify associations among corresponding information.

In summary, the design of Experiment 2 matched the manipulations of the previous experiment with a different lesson. The primary difference in this lesson was that the visual materials were animated rather than static. The two stage process model of learning from multimedia predicts that learning would still improve when presentation pace is slowed. However, especially for transfer, animated materials accompanied by narration might be at a disadvantage due to a reduction in contiguity, and animated materials accompanied by text might be at an advantage due to forcing learners to switch back and forth between media.

3.1 Method

3.1.1 Participants

Two hundred and twenty four participants who received credit for participation in a Psychology class were randomly assigned to each of 6 conditions. An analysis of previous research in the domain of multimedia learning was conducted to determine group sizes.[3] Participants were the same as the previous experiment: there were 127 male participants and 97 female participants. Their age ranged from 16 to 26 years old (M = 19.43); 77 were freshman, 65 were sophomores, 40 were juniors and 39 were seniors. Eighty-eight percent were native English speakers and the remainder all reported that they spoke English fluently.

3.1.2 Materials

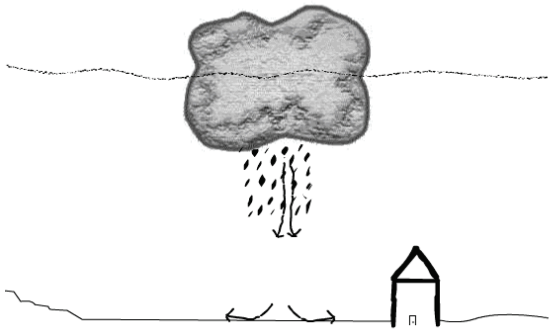
Lightning lesson. The lightning lesson was a multimedia presentation on the formation of lightning adapted from Mayer and Moreno (1998). The Flash presentation uses animation and verbal content to depict air moving from the ocean to the land, water vapor condensing to form a cloud, the rising of the cloud beyond the freezing level, the formation of crystals in the cloud, the movement of updrafts and downdrafts, the building of electrical charges within the cloud, the division of positive and negative charges, the

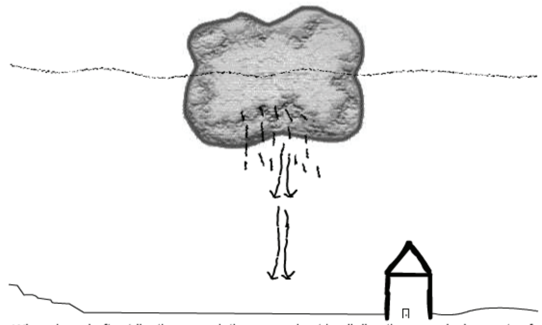
[3] G-Power software was used to compute the necessary sample size to detect a significant difference in a global comparison and a special comparison with numerator df = 2 (to test effects of manipulating of study time). The parameters used were and alpha = .05, power = .95, groups = 6 (because the planned experiment has six within subjects manipulations). A conservative eta of .33 was used to match that of the power analysis in Experiment 1 and based on the fact that this eta was less than effect sizes reported by Mayer (2001) related to the modality effect using the present materials. The analysis yielded a recommendation of 192 participants for the global comparison and 175 participants for the special comparison.

traveling of a negative stepped leader from the cloud to the ground, the traveling of a positive stepped leader from the ground to the cloud, the negative charges following the path to the ground, the meeting of the negative leader with the positive leader, and the positive charges following the path towards the cloud.

The animation was broken down into 16 sections, each containing 3-5 sentences. The multimodal version included concurrent narration describing each of the major events in segments that last between 3 and 9 seconds spoken at a slow rate by a female voice. Each segment began immediately after the previous section ended, with the beginning of an animated event and corresponding narration consistent across conditions. The unimodal version included a concurrent text (using the same words as contained in narrations) displayed on the screen while the animation segment played (see **Figure 8**). Besides the default length conditions (105 seconds; i.e., the narration time), there were two conditions with longer presentation times; one was 1.5 times the narration length (155 seconds) and the other was twice the narration length (210 seconds). In the unimodal cases the text was displayed for the entire segment. In the multimodal cases, the narration played at its default speed to preserve the clarity of the spoken verbal materials, however narrations were begun at the beginning of each segment to preserve synchronization. For instance in the longest multimodal condition, each animated segment was accompanied by 3-5 seconds of narration and 3-5 seconds of silence, but the corresponding words and pictures were always presented together.

Participants viewed this lesson immediately after completing the post lesson exam from Experiment 1. They were randomly assigned a condition In Experiment 2 with no regard to the condition they were assigned to in Experiment 1.





a. b.

Figure 8: Example screen shots from a) narration and b) text conditions of Experiment 2.

Individual differences measure. Before Experiment 1, the working memory capacity of participants was assessed using Automated O-Span.

Subjective measures. As with Experiment 1, participants completed a NASA TLX (Hart & Staveland, 1988) regarding the lesson. Participants also answered Likert-scale questions regarding the speed of the lesson, how they felt they performed on the post test, and their ability to identify how the materials correspond.

General learning test. Both before and after the lesson, participants were asked to write an explanation of how lightning works. Responses are interpreted to identify the number of idea units remembered out of 8. This test assesses both prior knowledge and general learning. See Appendix B.

Recall tests. Memory questions addressing specific details from the lesson about the heart and circulatory system were administered only after the learning phase. These questions assessed knowledge of facts and their basic relationships and were therefore appropriate to assess verbal learning outcomes (Gagne, 1972; Gagne et al., 2005). Questions were developed to include details that most participants would be unlikely to know or to be able to guess without having viewed the lesson. Five text-specific memory questions (see Appendix B) were answerable for a maximum score of 5 points.

Picture-specific memory questions were administered only after the learning phase. These questions presented four frames from the animation along with instructions to circle certain objects from the lesson (see Appendix B). Four visually related memory questions were answerable for a maximum score of 8 points.

Transfer tests. Inference questions related to the lesson were administered only after the learning phase. These questions assessed understanding of rules and principles inherent to the scientific system and relationships among parts of the system and were therefore appropriate to assess intellectual skill development (Gagne, 1972; Gagne et al., 2005). Some questions were picture-specific, some on text-specific. Three picture-specific transfer questions were answerable for a maximum score of 6 points and three text-specific transfer questions were answerable for a maximum score of 8 points (see Appendix B).

3.1.3 Procedure

Participants from Experiment 1 were also used for Experiment 2. After completing the pre-test, lesson, and post-test for Experiment 1, computer instructions directed the participants to continue onto Experiment 2 by completing the pre-test and then viewing the lightning lesson. After the lesson, participants were instructed to complete the post-tests.

The study design was a 2 (modality of verbal information) x 3 (study time) between subjects factorial design with each participant tested immediately following the lesson (immediate test). Additionally, they were tested after a 7 to 9 day retention interval (delayed test). Modality of verbal information was manipulated to be either text or narration. Study time was system paced for all conditions. However, the pace of presentation was manipulated under three conditions: timed to narration length (standard condition), 1.5 times narration length (150% condition), or double narration length (200% condition). Dependent measures accounted for general learning improvement (from pre-test to post-test), text and visually related verbal information learning outcomes, and text and visually related intellectual skill development outcomes.

The lighting lesson was administered in the second hour of testing, after participants had completed Experiment 1. Before Experiment 1, participants read and signed an experimental consent form, were briefed, informed that the six people who did best on the post lesson test would be awarded $10 and completed the Automatic O-Span. Following the lightning lesson, participants completed a paper and pencil test.

After the participant completed the experiment, a return appointment was scheduled for the delayed tests. When the participant returned, they were seated at a desk with instructions to complete the tests and allowed to leave when they had done so.

Coding data. Data were scored by three coders who were blind to treatment condition. Two coders rated each data point to check for consistency as necessary. To determine the necessity for a second rater, one primary coder rated about one third of the total data (data from 75 participants) and two secondary coders rated half those data (about 40 each). An individual question that was shown to have high inter-rater reliability (an individual question that had over 90% agreement: 7 or less discrepancies among the 75 sets of data) was deemed to no longer need to be rated by two separate coders. The three coders then met to clarify and discuss the questions with less than 90%

accuracy. Following this meeting, the primary coder again rated about one third of the total data (data from 70 participants) and two secondary coders rated half those data (about 35 each). Again, an individual question that was shown to have high inter-rater reliability (an individual question that had over 90% agreement: 7 or less discrepancies among the 70 sets of data) was deemed to no longer need to be rated by two separate coders. The three coders met to clarify and discuss questions with less than 90% accuracy. Following this meeting, the primary coder again rated the remainder of the data (data from 79 participants) and two secondary coders rated half those data (about 40 each). The three coders met again to discuss discrepancies.

3.2 Results and Discussion

As with Experiment 1, the lightning test was assessed for internal consistency reliability using Cronbach's alpha. The lightning test had a reliability of .70, so no items were deleted. Before analysis, outliers were deleted. As with Experiment 1, the removed outliers were those subjects whose total raw score (summed scores) for the lightning lesson was more than 2 standard deviations from the mean. Twenty-five subjects were removed from analysis by this method.

3.2.1 Learning

Manipulation check. The lesson, in all configurations, was successful in teaching learners about the formation of lightning: dependent sample t-tests showed reliable improvement in performance on the general knowledge test from pre- to post-test for all conditions (see Table 8). Correlations are weak in these comparisons because most participants did not know much about the formation of lightning before the lesson. This caused a restriction of range and an inability to rank order a large proportion of the participants based on pre-test scores. Nonetheless, it seems that most participants did learn about the formation lightning from the lesson.

Table 8: Mean performance for Pre- to Post-learning phase and t-test of differences for Lightning general knowledge test performance as a function of verbal presentation modality and presentation pace.

Modality	**Presentation Pace**	***N***	**Pre-Test (*SD*)**	**Post-Test (*SD*)**	**Correlation**	**t value**
Text						
	Standard	36	0.28 (0.57)	5.92 (1.82)	.023	17.810*
	150%	33	0.21 (0.54)	5.88 (1.49)	.109	21.286*
	200%	32	0.25 (0.51)	6.56 (4.41)	.202	25.482*
Narration						
	Standard	35	0.11 (0.40)	5.91 (1.27)	.192	27.329*
	150%	32	0.72 (1.37)	6.30 (1.12)	.032	17.057*
	200%	31	0.42 (0.76)	6.23 (1.54)	.002	18.324*

Note. Max score = 8. SD = standard deviation. * = p < .05

Overall learning. To assess the experimental hypotheses (i.e., assess the influences of verbal presentation modality and presentation pace on immediate and delayed learning) between subjects univariate ANOVAS were conducted with modality of verbal presentation (audio vs. narration) and presentation pace (three levels) as fixed factors. Separate univariate ANOVAS were conducted to avoid the pitfalls of using MANOVA with multiple dependent variables that might not be independently sampled. Four separate analyses, one with each test performance measure (immediate recall, immediate transfer, delayed recall, and delayed transfer) as dependent factors were performed. See Figure 7 for a representation of these data. The analysis of performance on immediate recall revealed a statistically significant main effect of presentation pace with a medium effect size, $F(2,189) = 3.35$, $p = .04$, $\eta_p^2 = .03$, no main effect for verbal presentation modality, $F(1, 189) = 1.12$, $p = .29$, and no interaction between the two, $F(2,186) = 1.59$, $p = .77$. See The analysis of performance on immediate transfer revealed a significant main effect of presentation pace, $F(2, 189) = 5.43$, $p = .04$, $\eta_p^2 = .03$, no main effect for verbal presentation modality, $F(1, 189) = 1.49$, $p = .22$, and no interaction between the two, $F(2, 189) < 1$. The analysis of performance on delayed recall revealed a marginal statistically significant main effect of presentation pace, $F(2,189) = 2.68$, $p = .07$, $\eta_p^2 = .03$, no main effect for verbal presentation modality, $F(1, 189) = 2.22$, $p = .14$, and no interaction between the two, $F(2,186) < 1$. The analysis of performance on delayed transfer revealed a statistically significant main effect of presentation pace, $F(2,189) = 4.97$, $p < .01$, $\eta_p^2 = .05$, no main effect for verbal presentation modality, $F(1, 189) < 1$, and no interaction between the two, $F(2,186) = = 2.28$, $p < .11$.

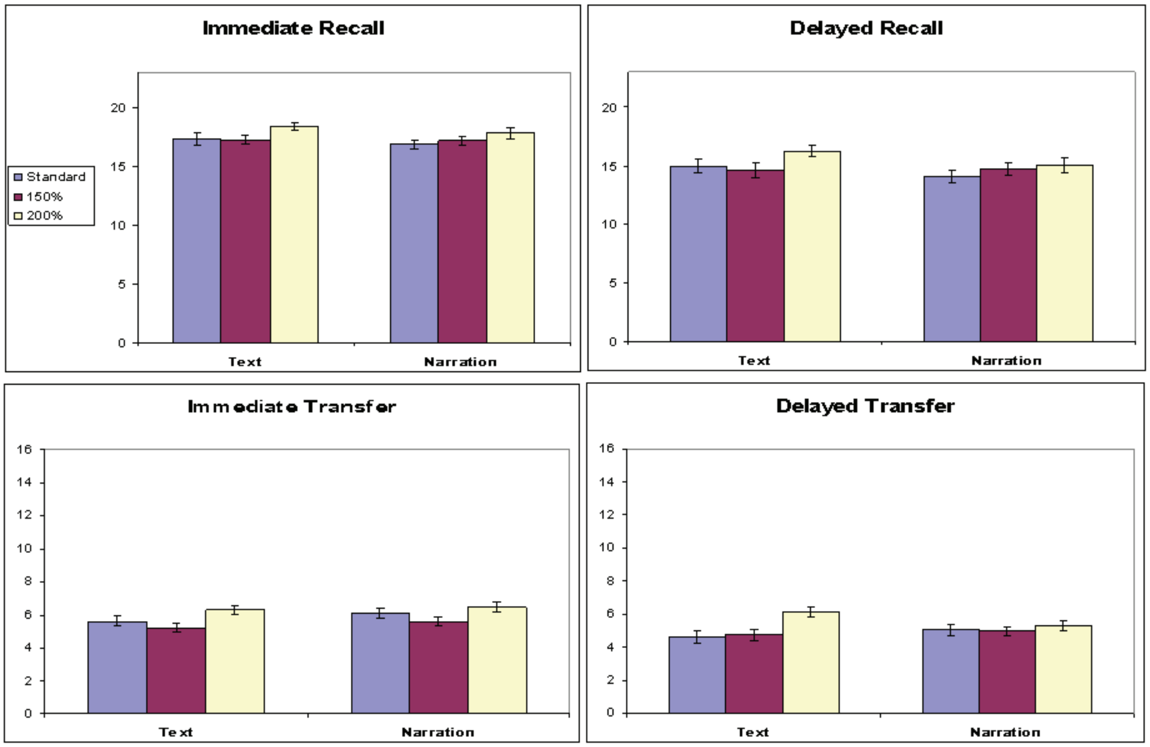


Figure 9: Chart depicting immediate recall performance following the heart lesson.

In summary, these analyses revealed significant effects with medium effect sizes of presentation pace on several individual measures of learning. There were significant effects of presentation pace (but not modality) on immediate recall, immediate transfer, and delayed transfer, as well as a marginal effect on delayed recall. To explore which group mean differences (see Table 9 for group means) were contributing to these effects, post-hoc LSD comparisons between levels of presentation pace collapsed over presentation modality were examined. As Table 10 indicates, these comparisons showed that, for every test type, group mean performance following the 200% pace condition was better than group mean performance following at least one other condition (performance following the standard pace condition, $p = .02$, and the 150% pace condition, $p = .05$, on immediate recall; performance following the 150% pace length condition, $p < .01$, on immediate transfer; performance following the standard pace condition, $p = .05$, on delayed recall; performance following the standard pace condition , $p < .01$, and the 150% pace condition, $p = .01$, on delayed transfer).

Table 9: Mean performance on each test administered immediately (immediate test) and 7 to 9 days after (delayed test) after the lightning lesson as a function of verbal presentation modality and presentation pace.

			Immediate Test		**Delayed Test**	
Modality	**Presentation Pace**	***N***	**Recall (*SD*)**	**Transfer (*SD*)**	**Recall (*SD*)**	**Transfer (*SD*)**
Text						
	Standard	35	17.29 (3.22)	5.61 (1.96)	14.94 (3.33)	4.64 (2.17)
	150%	31	17.23 (2.16)	5.19 (1.61)	14.55 (3.54)	4.76 (1.96)
	200%	31	18.39 (2.03)	6.29 (1.58)	16.26 (2.97)	6.16 (1.88)
Narration						
	Standard	34	16.82 (2.34)	6.06 (4.58)	14.06 (3.23)	5.03 (2.01)
	150%	31	17.19 (2.18)	5.58 (1.56)	14.71 (2.84)	4.98 (1.49)
	200%	31	17.77 (2.62)	6.47 (1.89)	15.03 (3.63)	5.29 (1.56)

Note: Maximum recall score = 23; Maximum transfer score = 16. SD = standard deviation.

Table 10: Mean performance on each test administered immediately (immediate test) and 7 to 9 days after (delayed test) after the lightening lesson in each presentation pace condition collapsed across presentation modality.

		Immediate Test		**Delayed Test**	
Presentation Pace	***N***	**Recall (*SD*)**	**Transfer (*SD*)**	**Recall (*SD*)**	**Transfer (*SD*)**
Standard	80	16.26 (3.63)*	5.59 (2.09)	14.04 (3.71)*	4.68 (2.21)*
150%	72	16.78 (2.71)*	5.31 (1.69)*	14.53 (3.32)	4.80 (1.89)*
200%	70	17.69 (2.93)	6.31 (1.86)	15.39 (3.42)	5.59 (1.84)

Note: Maximum recall score = 23; Maximum transfer score = 16. SD = standard deviation. * = worse than performance on 200% pace condition within modality and test type.

Presentation pace. These results confirm the hypothesis that more time to examine instructions allows learners to perform processes important to learning. In addition, the findings in Experiment 2 demonstrate that this effect may be extended to more than just measures of immediate recall. According to the two stage process model of learning from multimedia, transfer learning from animated materials accompanied by narration was expected to be hindered due to a reduction in contiguity, and transfer learning from animated materials accompanied by text was expected to be aided due to forcing learners to switch back and forth between media. To further explore this *a priori* hypothesis, I performed a one-way ANOVA with presentation pace as the independent factor and all learning performance measures as dependent factors. As with the previous experiment, there was a significant linear effect of slowing presentation pace on improving performance on immediate recall, $F(2, 196) = 5.67$, $MSE = 6.04$, $p = .02$. In addition, the present experiment also showed a significant linear effect of slowing presentation pace on improving immediate transfer, $F(2, 196) = 5.15$, $MSE = 2.87$, $p = .04$, delayed recall, $F(2, 196) = 2.35$, $MSE = 9.37$, $p = .03$, and delayed transfer, $F(2, 196) = 4.63$, , $MSE = 3.50$, $p < .01$ (see Table 10).

The fact that performance was consistently better when presentation pace was slowed was somewhat surprising. Research suggests that presenting corresponding words and pictures simultaneously rather than successively can foster retention (Mayer et al., 1999). As animated multimodal presentations were slowed in this experiment, the presentation contiguity of animated visual events and corresponding narrations were reduced. It was expected that the slow paced multimodal presentations in this experiment would lead to relatively worse learning (compared to fast paced multimodal presentation) due to a reduction in the likelihood that corresponding verbal and visual materials were presented at the same time when animation pace was slow (and the narrations were presented at a constant pace). This was expected to have a negative effect on the second process of learning from multimedia. In contrast, when animated unimodal presentations were slowed to provide more study time, the presentation contiguity of complementary information in the animated visual medium and text was not affected. Therefore, it was expected that the second process of learning from multimedia would be facilitated by text presentation because it would force learners to switch back and forth between text and animation, facilitating the identification how the information in the two media are related. This would be a benefit not provided, possibly even hindered, by multimedia presentation.

To examine more directly these hypothesized differential effects of presentation pace on learning from unimodal and multimodal versions of the animated lightning lesson, two separate follow up one-way ANOVAS were performed, one including the group who received audio verbal materials and one including the group that received textual materials. As presentation pace was slowed, there were marginally significant linear patterns of improving learning performance among text conditions -- immediate recall, $F(2, 98) = 3.20$, $MSE = 6.45$, $p = .08$, immediate transfer, $F(2, 98) = 2.73$, $MSE = 2.93$, $p = .10$, delayed recall, $F(2, 95) = 3.16$, $MSE = 10.85$, $p = .08$, and delayed transfer, $F(2, 95) = 10.10$, $MSE = 4.03$, $p < .01$ -- but no such pattern among narration conditions -- immediate recall, $F(2, 95) = 2.430$, $p = .12$, $MSE = 5.69$, immediate transfer, $F(2, 95) = 1.42$, $p = .24$, $MSE = 2.86$, delayed recall, $F(2, 94) = 1.58$, $p = .21$, $MSE = 10.47$, and delayed transfer, $F(2, 94) < 1$.

Though the inferential power of these data is weak, the consistency across all learning measurements suggests that slowing the presentation pace of instructional materials facilitated learning performance following animated lessons with text, but not animated lessons with narration. These results provide some additional support to the results of the previous study. Using post-hoc LSD comparisons for further examination of mean differences within text conditions shows that there were reliable pairwise differences demonstrating improved performance for each learning measure following the slowest presentation pace compared to one or both of the other pace conditions (e.g., immediate recall $p = .04$, immediate transfer $p = .01$, delayed recall $p = .03$, and delayed transfer $p < .01$). These findings extend the finding in Experiment 1 that the slowest text presentations led to the best performance on learning measures. These findings also suggest that the effect of slowing presentation of instructions using text is different from the effects of slowing presentation of instructions using narrations. It was predicted that this differential effect of pace on learning would cause a reduction or reversal of the modality effect. This pattern was expected to explain why there is typically no modality effect when presentation pace is not matched to the time it takes to play narrations (Ginns, 2005) and the results of previous studies reviewed in this paper that have compared the modality effect across conditions of varying presentation pace (Harskamp et al., 2007; Tabbers, 2002).

Modality effect. Due to the differential effects of presentation pace on learning from presentations using verbal materials with text versus narration to accompany animated materials, it was expected that the modality effect would be reduced and possibly reversed as presentation pace was slowed. More specifically, it was predicted that there would be a reduction of the modality effect as a function of slowing presentation pace in performance on recall tests and a reversal of the modality effect as a function of slowing presentation pace on transfer tests. This might also explain why Tabbers (2002) found a reduction and reversal of the modality effect as a function of providing more study time on tests of transfer but not tests of retention. However, the MANOVA presented above indicated that manipulating verbal presentation modality produced no reliable differences on individual learning performance measures, despite past research demonstrating the contrary (Mayer & Moreno, 1998).

To further explore the modality effect at each level of presentation pace in the present experiment, I combined the recall and transfer tests test into a summed score for the immediate test and a summed score for the delayed test. This combination was done in an attempt to increase the predictive power of these data and reduce the likelihood that the absence of the modality effect was due to insufficient power. I then performed 3 independent sample t-tests (one for each level of presentation pace) with verbal presentation modality as a grouping factor and the summed scores as dependent variables. Among those who received standard paced presentations there were no reliable differences in performance between those who received text versus narrations on the immediate test $t(69) = 0.15$, $p = .88$, or the delayed test, $t(68) = .51$, $p = .67$. This persistent lack of a modality effect is surprising. Dividing participants by major, OSpan, native English speakers, or year in school had no bearing on this effect. However, it is unlikely that this is due to insufficient power, because there was a reverse modality effect on delayed performance for slow paced presentations. That is, among those who received the 200% paced presentations, there was a reliable advantage on the delayed test for those who received the text presentation (mean score = 22.58) over narration presentation (mean score = 20.32), $t(61) = 2.03$, $p = .05$ even though there was no reliable differences in performance between those who received text versus narrations on the immediate test, $t(61) < .68$, $p = .53$. This is some support for past research that has suggested that the use of text to accompany animated materials encourages learning processes that are important to developing a detailed mental model and can be detected in delayed tests (Palmiter & Elkerton, 1993). According to the framework proposed here, these processes are related to the formation of associations between mental codes that lead to a better understanding of the workings of the system (i.e., intellectual skills) that is the subject of the lesson.

3.2.2 Subjective experience

Subjective measures of the learning experience were administered after the lessons. Table 11 shows mean responses on subjective ratings of workload taken by the

NASA-TLX. A univariate ANOVA examining the effect of the two presentation manipulations (modality of verbal presentation and presentation pace) on subjective ratings of cognitive load showed a main effect of presentation time, $F(2, 188) = 3.77$, $p = .03$, $\eta_p^2 = .04$ and a main effect of presentation mode, $F(1, 188) = 4.89$, $p = .03$, $\eta_p^2 = .03$. There was no interaction between the two manipulations, $F(2, 188) = 1.33$, $p = .27$. Post-hoc LSD pairwise comparisons showed that the slowest presentation caused significantly less subjective mental workload than system paced ($p = .01$) or medium paced ($p = .03$) conditions. These results parallel those of Experiment 1 (where slower paced presentations caused significantly less mental workload). Overall subjective mental workload was also lower for narration conditions compared to text conditions (see Table 11).

Table 11: NASA-TLX scores after the lightning lesson for each subscale and the summed total as a function of verbal presentation modality and presentation pace.

Modality	Presentation Pace	*N*	Total (SD)	Mental Demand (*SD*)	Physical Demand (*SD*)	Temporal Demand (*SD*)	Perform-ance (*SD*)	Effort (*SD*)	Frustra-tion Level (*SD*)
Text									
	Standard	36	298.33	59.74	7.44	73.46	51.15	61.67	51.03
			(100.85)	(26.06)	(12.77)	(23.06)	(22.81)	(23.23)	(28.40)
	150%	30	281.33	52.26	11.00	59.35	48.71	54.50	40.32
			(89.43)	(22.09)	(16.10)	(26.51)	(20.57)	(23.72)	(29.12)
	200%	30	231.00	56.67	7.27	49.55	40.61	51.21	34.70
			(91.45)	(45)	(16.25)	(23.60)	(20.70)	(24.18)	(28.91)
Narration									
	Standard	35	241.43	45.26	9.74	52.50	44.47	45.13	37.11
			(117.13)	(32.15)	(16.64)	(31.19)	(20.30)	(27.81)	(27.57)
	150%	32	252.97	53.43	12.00	47.57	50.00	51.57	41.86
			(79.41)	(22.16)	(17.49)	(23.21)	(21.14)	(20.10)	(23.55)
	200%	31	227.10	50.31	11.09	42.03	46.41	44.38	36.72
			(69.55)	(23.96)	(15.54)	(31.21)	(22.44)	(22.10)	(24.28)

Note: Maximum score for each subscale = 100; Maximum score for total = 600. SD = standard deviation.

Table 12 shows average responses on subjective performance questions. From these responses, it appears that the animated instructions were more susceptible to negative experiences based on presentation pace compared to the static materials of the previous experiment. Nearly 2 out of 3 (145 out of 220 in all conditions) participants said that they would slow the presentation down if they could. Responses for the lesson in Experiment 1 were much lower on average and suggested that participants would not have slowed even the fastest presentations form that experiment. In Experiment 2, nearly

2/3 of participants who received the animated lighting lesson answered 4 or 5 (would make it slower) on the subjective question about whether they would change the rate of presentation; the average response on the question was above 3 (3 = 'it was fine') even for the slowest presentations. This is also surprising, and might reflect that fact that no changes to the speed of the narration meant that there were no changes in the perceived speed of the lesson and that the lesson was not slowed enough to be easy to learn from when verbal materials were presented via text.

Table 12: Mean responses on the lightning lesson related Likert questions as a function of verbal presentation modality and presentation pace.

Modality	Presentation Pace	*N*	Speed (*SD*)	Understanding (*SD*)	Related (*SD*)	Performance (*SD*)
Text						
	Standard	36	4.18 (0.79)	2.69 (0.92)	2.05 (1.08)	3.10 (1.00)
	150%	30	3.84 (0.63)	2.25 (0.95)	1.56 (0.80)	2.75 (1.08)
	200%	30	3.60 (0.78)	2.23 (0.84)	1.60 (0.78)	2.63 (0.97)
Narration						
	Standard	35	3.92 (0.67)	2.50 (0.86)	1.97 (1.05)	2.63 (1.05)
	150%	32	3.37 (0.88)	2.31 (0.68)	1.86 (.97)	2.80 (.80)
	200%	31	3.44 (0.80)	2.16 0(.85)	1.72 (.92)	2.75 (1.30)

Note: Each score is out of 5. SD = standard deviation. See Appendix D for items.

CHAPTER 4: EXPERIMENT 3

Experiments 1 and 2 explored the relationship between the presentation pace and verbal presentation modality of instructional materials with learning outcomes. Based on the findings, it appears that reducing presentation pace can foster learning from multimedia materials using both static and animated visual materials. However, both experiments seem to indicate that slowing presentation pace fosters learning from materials including text more than those including narrations. Experiment 3 was aimed at extending this understanding of learning outcomes and correlating them with a process measure taken while people learned from the materials used in the previous studies. In this experiment I used an eye-tracking technique to measure the movements of learners' eyes as they observed the heart lesson and lightning lessons. Eye movements were recorded to gain insight into how learners interact with materials that lead to better or worse learning outcomes. This was done to extend the understanding of the two-stage process model of learning from multimedia by connecting outcomes to processes of learning by directly observing behavior that might be related to those processes. That is, the present experiment examined how learners approach and interact with multimedia instructional materials of different configurations known to have differential effects on learning performance.

Eye-tracking was used because it is less disruptive to learning (in terms of interrupting the learner) than think aloud protocols (cf., Butcher, 2006). Primary variables of interest related to learning from visual and verbal materials were length of time learners looked at diagrams, time reading textual materials, activities while listening to audio narrations (e.g., location of eye gaze, interaction with control widgets), and pattern of switching gaze between media with verbal and diagrammatic content. These measures were taken to confirm the suggestion that learning from multimedia is improved by allowing learners ample time to form mental codes to represent the to-be-learned information in both media, and the ability to identify how the information contained in two separate media correspond.

It is hypothesized that the first stage of learning proposed in this framework, the process of creating mental codes to represent the to-be-learned information, correlates with the ability to examine relevant parts of instructional materials. To test this hypothesis, patterns of eye movement were analyzed to measure the correlation of fixation time with trends in learning. If the formation of mental codes is dependent on fixation times (i.e., dwell time), one might expect to observe a significant correlation between the two factors. As we saw in Experiments 1 and 2, slowing the pace of presentation improves learning outcomes and, as hypothesized, these improvements are more significant for text presentations. In this experiment eye-movements were used to investigate whether changes in dwell time are more influential on learning as text presentation is slowed compared to changes as narration presentations are slowed. In addition, the amount of time spent fixating on picture information was compared between text and narration conditions that lead to similar learning outcomes.

A second hypothesis of the current study is that the second stage of learning proposed in this framework, the formation of associations between these mental codes, correlates with the ability to identify how concepts are related. To test this hypothesis, patterns of eye movement were analyzed to determine if there exists a correlation of switching behavior with trends in learning. It is expected that switching eye-gaze back and forth between media is related to the development of associations among to-be-learned information contained within different media and is therefore related to intellectual skill development outcome measures. Like with dwell time, I will explore eye-movement behavior to see if changes in dwell time are more influential on learning as text presentation is slowed compared to changes as narration presentations are slowed. This measure was not applicable when examining narration conditions of the lightning lesson because there was no reason for learners to look at the text regions. However, it was usable when examining narration conditions of the heart lesson because there was a visual progress bar in the text region that learners could look at and use during learning.

Eye movement research on the integration of pictures and text has shown that in most cases learners first read (at least part of) the text and then switch to the picture to integrate the verbal and the pictorial information (Hegarty et al., 1992; Rayner et al., 2001). It is possible that presentations with text timed to match the length of corresponding narrations hinder the acquisition of mental codes representing the to-be-learned information because they allow little time for learners to examine diagrammatic materials after reading text. The addition of study time should reduce this negative effect, leading to more total time available to view diagrammatic (whether they are static or animated) materials. If this is the case, the total time that the eye is set on diagrammatic materials should be similar between conditions that led to similar verbal information learning (recall) performance in previous experiments, even when the presentation might vary in modality of verbal information or study time provided.

As people listen to a story or follow instructions, they quickly move their eyes to those elements in an array that are most closely related to the words currently heard (Cooper, 1974). This searching of diagrammatic materials "on-line," while listening to narrations might foster the development of associations between mental codes representing to-be-learned information. However, research has suggested that the difference in effectiveness between system-paced and learner-paced instructions with text does not seem to be related to an overall difference in fixation time (Tabbers, 2002). A presentation with text might be easier for a learner to adapt to their individual needs by allowing them to control how they fixate on pictures and text. Learners cannot review narrations but can use additional time to examine diagrammatic materials. One might expect that as more study time is provided during presentations with text, switching behavior is similar to patterns observed in presentations with narrations that led to similar intellectual skill development outcomes (transfer performance) in the previous experiments. That is, viewing patterns should be similar when intellectual skill development is similar, even when the presentation might vary in modality of verbal information and presentation pace.

4.1 Method

4.1.1 Participants

Forty participants received $20 for taking part in this study. The participants were divided among 4 conditions including static materials for the first phase of the experiment and 4 conditions including animated materials for the second phase of the experiment, so 10 viewed each lesson. Previous eye-tracking experiments examining how people scan and view text and pictures have included 12 participants total (Tabbers, 2002), 8 participants per group (Hegarty et al., 1992; Hegarty, Mayer, & Monk, 1995) or 12 participants per group (Rayner et al., 2001).

4.1.2 Apparatus

A RED III corneal reflection camera and iView X software package (SensoMotoric Instruments) were used to monitor eye position. Before the experiment, a 9-point calibration with corner correction was performed. For this calibration the software took one measurement while the participant was looking at dots in the four corners, four sides of the screen, and one in the center (i.e., 9 points). The software computed relative position of the pupil and corneal reflection for each measurement and used these to calculate the fixation coordinates during the trial. Eye fixation coordinates were translated to record eye movements relative to standardized regions of interest established separately for each lesson. The heart lesson was presented using a series of web pages and Internet Explorer. The pictures and text were presented in the same place on each page. Two regions of interest were defined, one that encompassed the diagram section and one that encompassed the text section (see **Figure 10**). The former was used to identify when participants were viewing the diagram, and the latter to identify when participants were reading the text. The lightning lesson was divided into five regions of interest. These regions divided the lesson into four areas corresponding to the major animation events that can be tied to different sections of the lesson and one for the text area (see **Figure 11**).

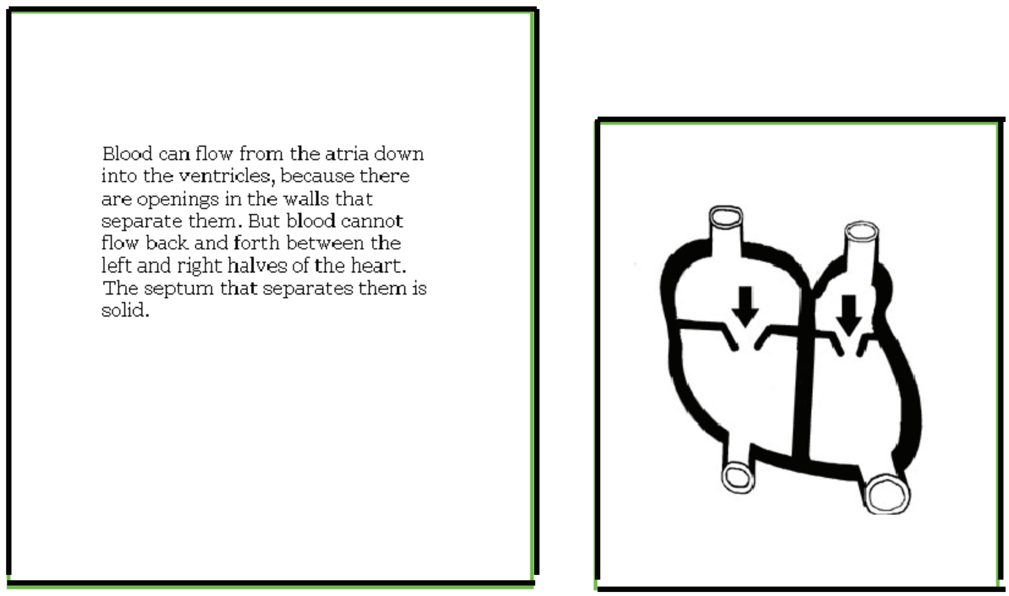


Figure 10: Regions of interests of the heart materials. The text region is on the left, the picture region is on the right.

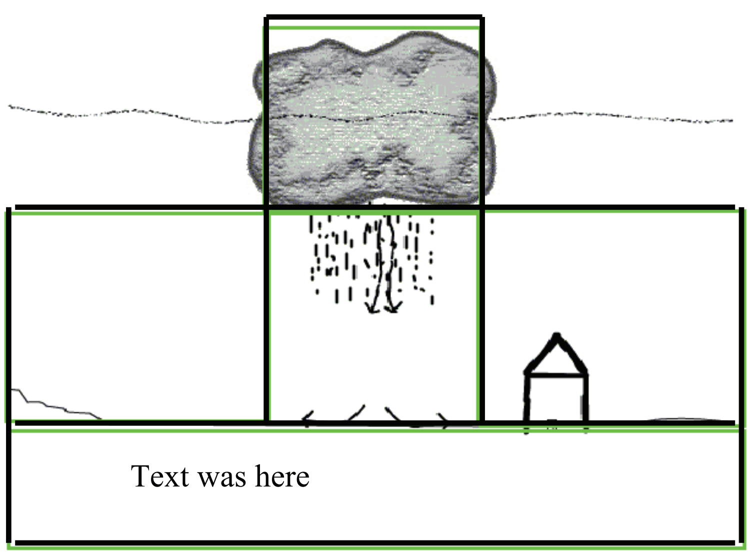


Figure 11: Regions of interest for the lightning materials. The text region is on the bottom (and has text in text conditions). Region one is the region on the left, Region 2 is the center region, Region 3 is on the right and Region 4 is on the top

4.1.3 Materials

The learning and test materials used in this study were the same learning and test materials used in Experiments 1 and 2. However, only four conditions from each experiment were used, so the design in the experiment was not fully crossed. Among the 8 conditions using static visual materials (heart lesson), I selected the system paced text and narration configurations, the text condition that provided double the study time, and the narration condition that included fifty percent more study time. The two system paced presentations were chosen to explore the absence of the traditional modality effect in Experiment 1. I expected to find differences in eye-movement behavior that corresponded to differences in learning. I chose the text condition that provided double the study time for the same reasons: to compare with system paced text conditions. Because there were no significant differences across conditions using narrations, I selected the narration condition that included fifty percent more study time. For instructions using animated visual materials (the lightning lesson) I selected the two system-paced and double-paced presentations to explore how changes in study time (the main factor that led to reliably different performance in Experiment 2) influenced learning.

4.1.4 Procedure

The study design was a 2 (modality of verbal information: visual-text, audio-narration) x 2 (standard versus slower presentation pace) between subjects factorial design. Gaze duration in each region of interest and number of transitions between text and diagram regions were the dependent measures. Learning measures (identical to the measures in Experiments 1 and 2) were also taken to validate the appropriateness of each trial in a cell and to check whether learning patterns are altered by eye-tracking procedures.

Students were tested one at a time. Upon arriving to the laboratory, participants read and signed an experimental consent form and informed that the four people who did best on each post lesson test would be awarded $25. They were then briefed that they needed to keep their head (with the help of a chin rest) still so that the camera could record the movement of their eye and told that they will be viewing two multimedia lessons. Participants then took the Ospan and heart pre-test. Next, they were seated at the computer workstation with the eye-tracker. They rested their chin on the chin rest and the calibration procedure began. Following calibration, they viewed the heart lesson. When the lesson concluded, participants removed themselves from the eye-tracking device and completed the paper and pencil heart learning measures and the lightning pre-test. After completing the measures they went back to the chin rest and the system was recalibrated. Following the second calibration, they viewed the lightning lesson. When the lesson concluded, participants completed the paper and pencil lightning learning measures.

4.2 Results and Discussion

To asses the efficacy of the data gathered, the learning tests were assessed for internal consistency reliability using Cronbach's alpha. The heart test had a reliability of 0.793 and the lightning test had a reliability of 0.56. No items were deleted from either test to remain consistent with the previous study and because learning was not the focus of this experiment. Before conducting any data analyses, outliers were removed for each lesson separately in the same way that outliers were removed in Experiments 1 and 2. The removed outliers were those participants whose total raw score (summed scores) was more than 2 standard deviations from the mean. This resulted in three participants being removed from the analysis of the heart lesson and 4 subjects being removed from the analysis of the lightning lesson.

4.2.1 Overall Learning

Both lessons, in all configurations, were successful in teaching learners about their subject, with dependent sample t-tests showing reliable improvement in performance from pre- to post-test $p < .01$ for all conditions of both lessons. To assess whether learning patterns were similar to the previous experiments, I performed an analyses of learning for each lesson that were parallel to the analyses in Experiments 1 and 2: separate univariate ANOVAs for each performance measure (recall and transfer) as dependent factors with modality (audio vs. narration) and study time (standard vs. slower) as fixed factors. See **Table 13** for cell means and standard deviations for both lessons.

As in Experiment 1, presentation pace had a significant effect on performance. However, in Experiment 3, the effect was evident on both the recall test, $F(1, 31) = 9.41$, $p < .01$, $\eta_p^2 = .23$, and the transfer test, $F(1, 31) = 4.28$, $p = .04$, $\eta_p^2 = .13$. No other significant effects of (verbal presentation modality or an interaction) were detected. These results differ from Experiment 1 in that the previous experiment detected differences in performance on the recall test but not the transfer test. Results related to performance on tests after the lightning lesson were also slightly different than those of the previous experiment. In Experiment 3, there were also significant effects of

Table 13: Mean performance on each test administered immediately after each of the heart and lightning lessons as a function of verbal presentation modality and presentation pace.

		Heart Lesson			Lightning Lesson		
Modality	**Presentation Pace**	***N***	**Recall (*SD*)**	**Transfer (*SD*)**	***N***	**Recall (*SD*)**	**Transfer (*SD*)**
Text							
	Standard	7	17.36 (4.68)	8.86 (3.24)	8	15.38 (2.56)	3.63 (3.29)
	Slower	10	19.80 (4.02)	9.10 (2.02)	7	18.29 (2.29)	6.71 (1.11)
Narration							
	Standard	9	16.50 (3.32)	6.44 (2.79)	9	17.90 (1.29)	5.70 (1.49)
	Slower	8	20.81 (3.83)	10.25 (3.85)	10	17.80 (1.55)	5.50 (1.65)

Note: Lightning maximum recall score = 23; maximum transfer score = 16. Heart maximum recall score = 29; maximum transfer score = 20. SD = standard deviation.

the interaction of presentation pace with verbal presentation modality on both recall, $F(1, 29) = 5.22$, p = .03, $\eta_p^2 = .15$ and transfer, $F(1, 29) = 5.97$, $p = .02$, $\eta_p^2 = .17$. Main effects were also significant for presentation pace on both recall, $F(1, 30) = 4.540$, $p = .04$, $\eta_p^2 = .13$ and transfer, $F(1, 30) = 4.37$, $p = .04$, $\eta_p^2 = .13$. Examining the means in **Table 13** suggests that the effects of pace on both recall and transfer were stronger as the pace of unimodal presentations was slowed than when the pace of multimodal presentations was slowed.

The results of Experiment 3 are similar to, but do not equal, the results of Experiments 1 and 2. For the heart lesson, the present data showed significant effects of presentation pace on immediate recall results as in Experiment 1; in addition the present results also showed significant effects of presentation pace on immediate transfer. For the lightning lesson, both Experiment 2 and Experiment 3 showed significant effects of presentation pace, but Experiment 3 showed a significant interaction of presentation pace and verbal presentation modality. For all three experiments, the effect of slowing pace on learning from text seems to be slightly greater than the effect of slowing pace on learning from narrations.

The primary difference between the present testing environment and the testing environment from the previous studies was the pressure for participants to concentrate on the lesson and test. This may have led to a more valid test in Experiment 3 compared to Experiments 1 and 2. Due to the eye tracking procedure, each participant was given more individual attention, possibly giving the participants an impression of importance or made them try harder because someone was watching them. In addition, the testing conditions made it so that participants were less likely to stop paying attention to the materials if they were too slow or boring, probably improving the effects of the pace manipulation. Participants in this final study were also being compensated monetarily for their time in the eye-tracker and received a greater monetary reward if they did well on the tests of learning. It is also possible that this pressure forced students to pay attention to the lesson and prompted more guesses and better recall of ideas related to transfer questions following the heart lesson, earning more points on this test and leading to detectable differences in test performance. In addition, there may be less variance due to being tired or losing concentration for quick animated materials under these individual testing conditions compared to group testing of the previous experiments. These environmental conditions are all more similar to real learning environments, where students feel more pressure to do well on tests. This may have caused the small differences (primarily effects on transfer following the heart lesson) in outcomes for this set of data. Future studies can explore the influence of these factors on learning, and the results herein should be interpreted while taking this testing condition in consideration.

4.2.2 Measuring eye-movement behavior

Dwell time was measured on a gross level for each participant. I-View Analysis Software (SensoMotoric Instruments) was used to measure the total time that the eye was directed toward each region of interest. In addition, total unmeasured time (lost data attributed to saccades, blinks, an inability to obtain measurements of the pupil or corneal reflection, etc.) and total time for the lesson were recorded. From these data, measured time (total time minus unmeasured time) and time looking at parts of the screen other than the regions of interest (neither region; i.e., measured time minus the sum of all regions of interest) were computed. Percentages of measured time for each region and neither region were computed (raw measurements for each divided by the measured time). These account for 100% of the measured data. Finally, percentage of undefined time was computed (raw unmeasured time divided by total time). See Table 14 for percentages on the heart lesson and Table 15 for the lighting lesson.

Table 14: Percent of time spent viewing regions of the heart lesson spent viewing the text and picture regions of the heart lesson, other parts of the display (neither region), and time that no measurement was taken (undefined) as a function of verbal presentation modality and presentation pace.

			Text Region (% of measured time)		Picture Region (% of measured time)		Neither Region (% of measured time)		Undefined (% of total time)	
Modality	**Presentation Pace**	***N***	***Min - Max***	***M* (*SD*)**	***Min - Max***	***M* (*SD*)**	***Min - Max***	***M* (*SD*)**	***Min - Max***	***M* (*SD*)**
Text										
	Standard	7	.67 - .92	.83 (.087)	.07 - .28	.14 (.075)	.01 - .06	.03 (.017)	.01 - .16	.06 (.046)
	Slower	10	.67 - .78	.74 (.040)	.13 - .25	.21 (.039)	.02 - .12	.06 (.036)	.01 - .15	.08 (.055)
Narration										
	Standard	9	.10 - .37	.25 (.095)	.56 - .85	.66 (.102)	.01 - .27	.11 (.087)	.01 -.29	.12 (.089)
	Slower	8	.17 - .36	.26 (.068)	.52 - .78	.62 (.075)	.01 - .06	.12 (.090)	.04 .36	.14 (.097)

Note: M = Mean; SD = standard deviation. Undefined time was not recorded, so the proportion of time viewing all regions sums to 100% of recorded time.

Table 15: Mean percent of time spent viewing each region of the lightning lesson, other parts of the display (no region), and time that no measurement was taken (undefined) as a function of verbal presentation modality and presentation pace.

Modality	**Presentation Pace**	***N***	Text Region (***SD***)	Region 1 (***SD***)	Region 2 (***SD***)	Region 3 (***SD***)	Region 4 (***SD***)	No Region (***SD***)	Undefined (***SD***)
Text									
	Standard	4	.72 (.07)	.07 (.05)	.01 (.01)	.05 (.04)	.11 (.04)	.06 (.07)	.06 (.06)
	Slower	6	.56 (.10)	.09 (.01)	.02 (.01)	.08 (.04)	.18 (.07)	.06 (.05)	.05 (.05)
Narration									
	Standard	6	.03 (.02)	.31 (.14)	.04 (.02)	.18 (.03)	.39 (.07)	.07 (.04)	.06 (.04)
	Slower	6	.02(.02)	.22 (.15)	.03 (.01)	.17(.04)	.51 (.15)	.04 (.02)	.04 (.02)

Note: M = Mean; SD = standard deviation. Undefined time was not recorded, so the proportion of time viewing all regions sums to 100% of recorded time.

For the heart lesson, switches were computed on a slide by slide basis. This computation was performed only for slides that had pictures, a total of 29 slides. Measurements that did not fall within the two regions of interest were removed. One switch was computed each time the location of gaze was measured to be in the text region immediately following a measurement in the picture region or vice-versa. Data were summed to compute a total number of switches during the heart lesson. See Table 16 for means by condition. Switches were computed simultaneously over the entire lightning lesson for only text conditions; no switches were computed for narration conditions because there were no competing visual materials in the text region during this condition (unlike the heart lesson, which had the progress bar). One switch was computed each time the location of gaze was measured to be in the text region immediately following a measurement in the picture region or vice-versa. See Table 17 for means by condition.

Table 16: Number of times participants switched from text to picture regions or vice-versa while viewing the heart lesson as a function of verbal presentation modality and presentation pace.

Modality	**Presentation Pace**	***N***	***Min***	***Max***	***M***	***SD***
Text						
	Standard	5	11	107	70.80	36.540
	Slower	6	97	229	160.00	54.457
Narration						
	Standard	6	80	158	121.33	29.019
	Slower	5	120	200	173.20	32.668

Note: M = Mean; SD = standard deviation.

Table 17: Number of times participants switched from text to picture regions or vice-versa while viewing the lighting lesson with text as a function of presentation pace.

Modality	**Presentation Pace**	***N***	***Min***	***Max***	***M***	***SD***
Text						
	Standard	5	38	76	53.50	17.99
	Slower	6	36	96	67.00	24.71

Note: M = Mean; SD = standard deviation.

4.2.3 <u>Eye-movement behavior during heart lesson</u>

To assess the influences of verbal presentation modality and presentation pace on eye movement behavior during learning from the heart lesson, a multivariate analysis of variance (MANOVA) was performed with modality (audio vs. narration) and presentation pace (slow versus fast) as fixed factors, and two eye-movement behaviors (total fixation time within the picture region, and number of switches of gaze between regions) as dependent factors. Only one measure for fixation time (on either text or picture) could be used because MANOVA analysis is not robust when using two dependent measurements that are highly reliant upon each other (e.g., a participant cannot look at the picture if they are looking at the text). Fixation time on picture regions was chosen for the analysis because the model used to make predictions regarding learning in previous experiments suggests that standard text presentation leads to worse test performance than slow text because standard text presentation prevents learners from being able to form mental codes representing information contained in the visual materials. It was permissible to use percent of time viewing the picture and total switches together in this analysis because they were not significantly correlated with each other, $p = .15$ (i.e., not dependent upon each other).

In predicting general eye movement behavior measures, the multivariate test of differences between groups based on *presentation modality* using the Wilks Lambda criteria was statistically significant with a large effect size, $F(2, 17) = 88.29$, $p < .01$, $\eta_p^2 = .91$. The multivariate test of differences between groups based on *presentation pace* using the Wilks Lambda criteria was also statistically significant with a large effect size, $F(2, 17) = 9.41$, $p < .01$, $\eta_p^2 = .53$. There was no significant interaction, $F(2, 17) = 1.09$,

$p = .36$, $\eta_p^2 = .11$. This suggests that participants who received presentations with narrations reliably spent a considerably larger amount of time viewing the pictures (see Table 14 for cell means). This is not surprising because the competing visual stimulus in the narration condition (i.e., the progress bar) carried less information than the competing visual stimulus in text presentations (i.e., the text). Somewhat surprisingly, participants who received presentations with narrations also made considerably more switches between regions (see Table 16 for cell means). This is surprising because there was little informational utility to the learners in switching back and forth between the picture and the progress bar. However, this might indicate that participants in text conditions were unable to make switches because they were busy reading. The MANOVA results also suggest that participants in slower presentation pace conditions spent a considerably larger amount of time viewing the pictures and made considerably more switches between the text and picture region.

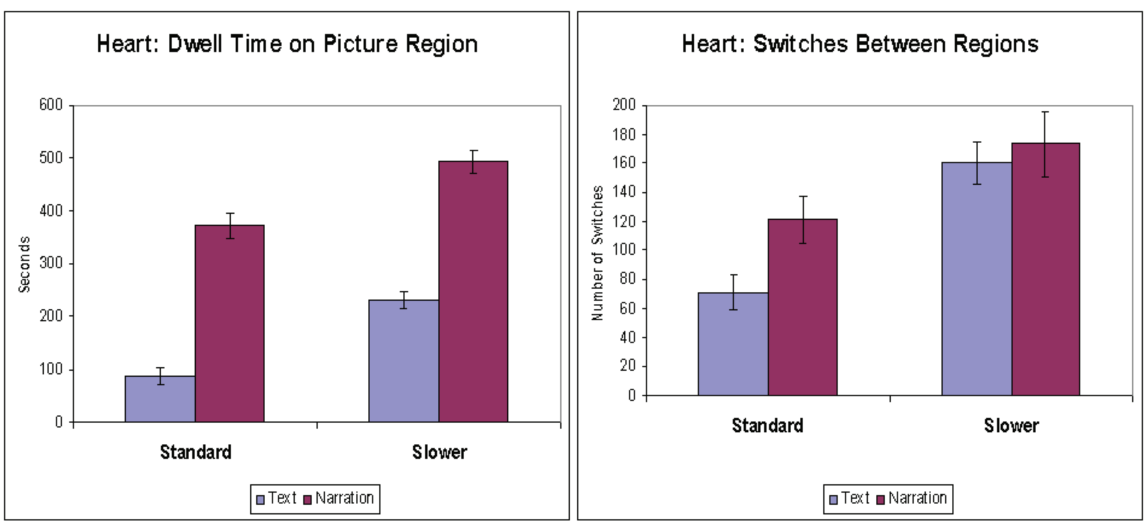


Figure 12: Eye-tracking data during heart lesson

In summary, these results confirm the hypothesis that learners viewing standard text presentations are likely to spend less time viewing pictures and make fewer switches between the text and picture region. See Figure 12 for a representation of these data. That is, the results suggest that learners in other conditions (compared to standard text presentations) are likely to spend more time viewing pictures and make more switches between the text and picture region. It might be that this lack of time spent viewing pictures while learning from multimedia hinders the first stage of the two-stage process model: the formation of mental codes to represent the information contained in the lesson (specifically the pictures). It might also be that this lack of switching back and forth between the text and pictures while learning from multimedia hinders the second stage of the two-stage process model: the formation of associations between these mental codes. These facts are further explored below in sections specific to each dependent measure.

Dwell time. It was hypothesized that the total dwell time on visual materials should be different between conditions in which verbal learning outcomes (i.e., recall test performance) are different, regardless of differences in the modality of verbal presentation or the presentation pace. In Experiment 1, separate one-way ANOVAs on presentation pace for the group who received narrations and the group that received text showed that recall test performance was reliably better between the standard paced text condition and the 200% paced text condition, but there were no reliable pairwise differences among narration conditions. To parallel this analysis, separate t-tests for the narration and text groups were performed to compare percent dwell time on pictures between slow versus fast paced conditions. This analysis of eye-movement behavior revealed a parallel pattern: percent dwell time on the picture region was significantly greater for the 200% paced text condition than the standard paced text condition, $t(14) = 2.20$, $p = .05$, mean difference = 7.00%, but there were no differences between narration conditions, $t(14) = 1.01$, $p < .33$ (see Table 14 for percentages). That is, the patterns in recall performance on the heart lesson (from Experiment 1) parallel those of the percent dwell time on picture regions in the present experiment. This is support for the hypothesis that the increases in learning due to reducing presentation pace are related to learners' ability to spend a more appropriate proportion of learning time viewing pictures. This enables the learners to form mental codes representing the information contained in the instructions.

In addition, it appears that the time looking at the pictures was productive in terms of improving performance on post lesson tests in text conditions. The correlation between time spent looking at the picture and overall performance on learning tests was significant for those in the text conditions, $df = 14$, $r = .58$, $p = .02$. However, the time spent looking at pictures did not predict performance for those in the narrations conditions, $df = 14$, $r = .02$, $p = .94$. This suggests that the time spent looking at pictures is important for learning from instructions presenting verbal materials via text, but perhaps that there is not practical difference in learning related to time spent looking at pictures when instructions present verbal materials via narration.

Examining the mean amount of time viewing picture materials shows that participants in standard paced text condition spent considerably less time (about 87 seconds) viewing the picture region compared to any other group (see Table 18). However, the time spent viewing the picture region in the 200% paced text condition was still less than either narration condition. It is likely that the participants in the slow paced text condition had plenty of time to view the materials and used extra time to re-read the text. Participants in the narration conditions did not have text to read and so may have continued to examine pictures even though there was a decreasing utility in such behavior. I had expected that participants in the slow narration condition would use this extra time to replay parts of the narration. However, not one participant - in any condition - even attempted to replay the narrations. Instead they appear to have looked around at other parts of the screen (percent of measured time spent outside the picture and text regions of the screen was significantly greater for narration presentations than text presentations, $t(33) = 2.26$, $p = .02$, mean difference = 6.0%) and were observed to sometimes even close their eyes as they listened to narrations (as suggested by reliably greater percent of undefined time for narration presentations than text presentations, $t(33) = 3.10$, $p < .01$, mean difference = 7.1%).

Table 18: Raw amount of time (in seconds) viewing text and picture regions of the heart lesson as a function of verbal presentation modality and presentation pace

			Text Region (seconds)		Picture Region (seconds)	
Modality	**Presentation Pace**	***N***	***Min - Max***	***M (SD)***	***Min - Max***	***M (SD)***
Text						
	Standard	7	412.25 – 585.70	513.21 (55.20)	40.64 – 171.85	87.67 (46.96)
	Slower	10	701.45 – 903.36	819.66 (67.98)	148.16 – 278.32	231.44 (45.72)
Narration						
	Standard	9	56.82 – 236.55	131.93 (56.98)	286.53 – 520.60	371.68 (76.47)
	Slower	8	116.23 – 331.26	215.67 (73.94)	116.23 – 331.26	492.81 (66.23)

Note: M = Mean; SD = standard deviation.

Switching between regions. It was hypothesized that increased occurrence of switching between text and picture regions would be different between conditions in which intellectual skill development outcomes (i.e., transfer performance) were different, regardless of differences in the modality of verbal presentation or the presentation pace. However, there were no reliable differences in transfer test performance detected in Experiment 1. Nonetheless, the number of switches improved greatly as presentation pace was slowed no matter the verbal presentation modality. It is possible that the ability to switch back and forth had little utility in the heart lesson. It seems likely that this behavior was done only because participants were forced by the procedures to observe the lesson (remaining in the chin rest during the entire presentation). Future studies might better address the issue of eye switching behavior in self-paced studies, where learners will be able to determine how often they need to switch between media themselves, or with more complex diagrams that require reference to the text materials for understanding.

4.2.4 Eye-movement behavior during lightning lesson

To assess the influences of verbal presentation modality and presentation pace on eye movement behavior during learning from the lightning lesson, the influence of these variables on viewing picture regions was assessed. The lightning lesson materials did not have a place holder for verbal media in narration conditions (i.e., there was no equivalent to the progress bar in the heart lesson). Therefore, switch behavior could not be included because there was no such measurement for narration conditions. This also presented a problem for comparing patterns of viewing animation. To equalize the measurement between conditions with different verbal presentation modalities, percent of time viewing the animation devoted to each region was assessed. See Table 19 for means. A multivariate ANOVA could not be performed with these dependent variables due to their dependence upon each other. Each was highly dependent upon at least one other: See Table 20 for a correlation matrix.

Table 19: Proportion of time looking at the lightning animation that was directed at each region of the heart animation as a function of verbal presentation modality and presentation pace

Modality	Presentation Pace	*N*	Region 1 ***M*** (***SD***)	Region 2 ***M*** (***SD***)	Region 3 ***M*** (***SD***)	Region 4 ***M*** (***SD***)
Text						
	Standard	4	.26 (.13)	.05 (.04)	.17 (.12)	.39 (.07)
	Slower	6	.22 (.08)	.06 (.02)	.18 (.05)	.41 (.13)
Narration						
	Standard	6	.32 (.15)	.04 (.02)	.19 (.03)	.40 (.07)
	Slower	6	.22 (.16)	.03 (.01)	.17 (.04)	.52 (.15)

Note: M = Mean; SD = standard deviation.

Table 20: Correlation of time looking each region of the lightning animation.

	Region 1	Region 2	Region 3	Region 4
Region 1	1	.181	-.209	-.709*
Region 2	.181	1	-.470*	-.374
Region 3	-.209	-.470*	1	-.113
Region 4	-.709*	-.374	-.113	1

Note: * p = < .05

Therefore, a series of univariate analyses of variance (ANOVA) were performed with modality (audio vs. narration) and presentation pace (slow versus fast) as fixed factors, and percent of time viewing each picture region during the animation as a dependent factor. The only significant effect detected (in all four analyses) was an effect of verbal presentation modality on viewing Region 2. Participants in text conditions were significantly more likely to spend more time viewing Region 2 that participants in the text conditions, $F(1, 18) = 4.56$, $p = .05$, $\eta_p^2 = .20$. However, this is likely due to the proximity of this region to the text region. Participants in the text condition may have looked at Region 2 as they scanned from an animation region to the text region. These effects suggest that neither modality of verbal presentation nor presentation pace had an effect on the proportion of time participants viewed each region of the animated materials.

Dwell time. As suggested by the unvariate ANOVAs reported above, there were no meaningful differences in viewing regions of the animation based on condition. Not surprisingly, participants in the text condition viewed the text region substantially more than participants in the narration conditions. Moreover, viewing of each region appeared to be timed to visual events in that region.

Switching between regions. There were no recordings of switch behavior for participants in the narration presentation of the animated lightning lesson. In addition, there were no significant differences based on presentation pace of text conditions. This suggests that people might not have paused in the middle of reading to look at the picture. Further exploration would be beneficial to more accurately record the timing of movements to the playing of animation to examine whether learners returned to text before words appear in slow paced but after in fast paced. This investigation would require a different experimental procedure and would also benefit from longer animations with variable segment lengths.

In summary, the results of Experiment 3 support the two-stage model of learning from multimedia and its predictions regarding how interactions with instructions (i.e., looking at media) of different configurations are related to learning outcomes. For instance, patterns of increasing percent dwell time on pictures in the heart lesson with reduced presentation pace, an interaction assumed to be necessary to perform the first stage of learning from multimedia, parallel the patterns of increasing recall performance in the heart lesson with reduced presentation pace, a learning outcome assumed to be a result of being able to perform first stage of learning from multimedia. These results explain the theoretical framework of the two stage model by mapping some interaction behaviors onto each stage. This is both a confirmation and explication of the model proposed as one way to understand how people learn from multimedia instructional materials. This detailed examination of learning from multimedia helps show why presentation pace has a larger influence on learning from instructions using text than it does on narrations (more appropriate proportions of viewing each medium and switching back and forth) and why the modality effect might not occur when presentations are not matched to the length of narrations (these interactions are not materially different when there is ample time to examine materials).

CHAPTER 5: GENERAL DISCUSSION

The goal of this research was to explore a new explanation for why people learn better when verbal materials that accompany diagrams are presented via narrations than when presented via text (Mayer & Moreno, 1998; Moreno & Mayer, 1999; Mousavi et al., 1995). Extant theories (CLT and CTML) attribute this *modality effect* to an expansion of working memory. However, recent research has questioned the appropriateness of this explanation on both empirical and theoretical grounds (Tabbers, 2002). The experiments described herein test the predictions of the expansion of working memory hypothesis against an alternative explanation for the modality effect: that using narrations allows learners to sense verbal and diagrammatic materials simultaneously, thereby helping learners identify relevant information in both the verbal and diagrammatic materials and how this information is related earlier in the course of learning. That is, dual mode presentation allows learners to form mental codes to represent to-be-learned information from both media simultaneously and also to more efficiently scan diagrams to determine how they are related to these verbal materials.

When instructions with animated or static diagrams are matched to the pace of presentations, using narrations is better suited to facilitating the two processes of learning from multimedia due to this advantage of narration presentation (quicker and more efficient learning). The two processes of learning from multimedia might also be facilitated when the presentation pace of instructions with text and static diagrams is slowed. In contrast, there are no gains to be made in learners' ability to perform the two processes of learning from multimedia by slowing the presentation pace of instructions with narrations and static diagrams. With animations, the second process of learning from multimedia (determining how the verbal and diagrammatic information is related) might be reduced when instructions using narrations are slowed because this might reduce the contiguity of the two media. In contrast, the second process of learning from multimedia might actually be encouraged when animated instructions use text are slowed because learners are forced to look back and forth between text and animation. This combination of factors might lead to a reverse modality effect when slowing the presentation pace of animated instructions.

Three experiments examining learning from qualitatively different instructional materials reported in this paper support these hypotheses. These results also confirm that predicting learning based on the expansion of working memory hypothesis can be inaccurate while predicting learning based on the two-stage model can be more accurate.

In Experiment 1, when accompanying diagrammatic materials were static, there was evidence of a reliable improvement in recall test performance (believed to be related to the formation of mental codes) when presentation pace was slowed for instructions using text, but there was no reliable improvement when presentation pace was slowed for instructions using narrations. This is predicted by the proposed model; the expansion of working memory hypothesis would predict that slowing presentation pace would improve learning regardless of verbal presentation modality. In Experiment 3, these learning patterns were shown to be identical to theoretically related patterns in eye-movement behaviors. That is, dwell time on pictures (believed to be related to the formation of mental codes) increased when presentation pace was slowed for instructions using text, but there was no reliable increase when presentation pace was slowed for instructions using narrations. This correspondence between measures, each of which are thought to be related to the formation of mental codes, is strong support for the first stage of the proposed model.

In Experiment 2, when accompanying diagrammatic materials were animated, the pattern of effects was largely the same. However, performance on the delayed test, which are often very sensitive measures of learning (Schmidt & Bjork, 1992), suggest that those who received longer (200%) text presentations actually learned better than those who received longer (200%) narration presentations. That is, it appears that animated materials with narration do not derive the same benefits of slowing presentation pace as animated materials with text. This is predicted by the proposed model; the expansion of working memory hypothesis would predict that slowing presentation pace would improve learning regardless of verbal presentation modality. Moreover, it supports the hypothesis that the use of text to accompany animated materials is better compared to narrations in terms of encouraging learning processes that are important to developing a detailed mental model (Palmiter & Elkerton, 1993). If these learning processes include determining how the information contained in verbal and diagrammatic information is related to each other, this is strong support for the second stage of the proposed model.

CHAPTER 6: CONCLUSIONS

The results from the three experiments reported here seem to support the two-stage-process model to a greater degree than they support the expansion of working memory hypothesis. Unfortunately, these results did not demonstrate changes in reliable differences between conditions of instructions using text and narrations. However, it might be that this is more a result of the experimental conditions than an inaccurate theory. In Experiment 1, I expected learners in the standard paced conditions to experience time pressure compared to conditions of slower pace. If this were the case, the relative efficiencies of learning from narrations would have been detected. However, participants in standard paced conditions of the heart lesson reported low temporal demand compared to participants in standard paced conditions of the lightning lesson. In addition, few participants expressed a desire to slow the presentation pace of the heart lesson. This suggests that the pace of the standard paced heart lesson was too slow to adequately test the modality effect. This seemed to be confirmed by eye-tracking results showing that participants in the fast paced text conditions were able to read and sometimes reread the text. In future experiments, I would use a more quickly paced narration to set the standard time. In addition, it might be useful to use a few levels of narration speed (e.g., words per minute) as a manipulation of presentation pace.

Based on this evidence, it seems that the levels of pace I chose may not have allowed me to accurately test the model in the exact way I planned. If presentation pace for multimedia instructional materials can be thought of as a continuum from lightning 'fast' to excruciatingly 'slow,' the portion of that continuum tested in this study might have begun closer to the 'slow' end than I hoped it would. This may explain why the traditional modality effect (at standard paced presentations) was weak in these studies. Especially in the heart lesson, there might not have actually been a condition with pacing equivalent to presentations used in previous studies that have demonstrated the modality effect. However, this study still found benefits to slowing the pace for text presentations, indicating that the benefits to slowing presentation pace of instructions with text might extend extremely far into the ‘slow’ end of the continuum.

While not a focus of this study, the potential for abenefit to an extremely slow presentation pace of instructions using text might indicate that instructional materials should be designed in ways that encourage learners to utilize study time to reexamine text. This might suggest that the ideal formatting for self-paced materials (such as the heart lesson) is to combine an entire lesson into one slide. I suggest this because learners who have a series of slides to study will likely go through each once as they study, and be less likely to review the materials in ways that are necessary to complete the two stage process of learning from multimedia. That is, they are less likely to utilize their time by reexamining materials as they were forced to under the 200% pace conditions of Experiment 1. However, if they have unlimited time to study and all of the information is in front of them, they might be more likely to reexamine the information that is available. This might make single slide presentation better for facilitating the two stage process of learning form multimedia. In contrast, Cognitive Load Theory and Cognitive Theory of Multimedia Learning would predict that because a single slide would be more complex than a series of simpler slides, the former would lead to higher demands on working memory and hinder learning. This is evident in their common guideline that instructions should be presented in smaller segments to reduce load (Mayer, 2001; Sweller, 1999).

Examining these predictions might be another way to test the model proposed in this paper against extant models of learning from multimedia.

As expected, learners in the slow paced text conditions reviewed instructional materials with the time they had. This seemed to increase performance on tests of learning. However, observations made during the eye-tracking study suggest that learners in the slow paced narration conditions did not review the materials. In fact, several participants in longer narration conditions literally closed their eyes towards the end of the eye-tracking session. It was also common for participants to pay a lot more attention to the progress bar for first few slides and then to ignore it in later slides. No participants tried to replay narrations and, despite instructions recommending they replay narrations if they have the chance, most participants took their hand off the mouse before the first narration was done, immediately committing to letting the lessons play on their own. Future research examining learning from slower paced presentations using narrations needs to provide a better way for learners to replay narrations or portions of narrations. This is an interesting design problem and I believe that this study demonstrates an evidence based need for design solutions.

In summary, the experiments in this paper, despite their limitations, support the hypothesis that one major advantage to learning from multimedia presentations using narrations compared to text is the efficiency with which learners can perceive multimodal presentations. Moreover, as predicted, when the disadvantage of a lack of time to perform both of the two stages of learning from multimedia materials using text is reduced by slowing presentation pace, learning was more comparable to learning from multimedia materials using narrations (of any pace). These results also suggest that an accurate understanding of the mechanism that causes the modality effect can predict when the use of narrations might not be helpful to learning. It is important to have a theoretical basis that can accurately predict the interaction of design factors and not a litany of guidelines with caveats. This is important because it is not practical to investigate all potential combinations of design factors empirically, especially as new educational technologies emerge. Accurate predictions are needed because the financial and other resources needed to change instructional design, such as converting a unimodal presentation (text and pictures) to a multimodal presentation (narration and pictures), can be enormous. In many situations, if there is no, or even a marginal, resulting benefit to learning, the conversion can be a waste of these resources. Further research into understanding other guidelines for the design of multimedia materials should help clarify circumstances in which these resources should be used on such technology.

APPENDIX A:

HEART LESSON

1. How many types of blood vessels are there? (relevant)
   1 point
   3 types

1b. Please name the different types of blood vessels. (relevant)
   3 points max
   1) Veins (veinules)
   2) Capillaries
   3) Arteries (arterioles)

2. How many chambers are there in the heart? (relevant)
   1 point
   Four chambers

2b. Name the chamber(s) of the heart. (relevant)
   4 points
   Left & right atrium
   Left & right ventricles

3. Blood returning from the body enters which chamber of the heart first? (relevant)
   1 point
   Right atrium

4. What is another name for the right atrioventricular valve?
   1 point
   Tricuspid valve

5. What is the protein which makes quick oxygen/carbon-dioxide transfer possible?
   1 point
   Hemoglobin

5b. How many molecules of oxygen can each such protein carry?
   1 point
   4 molecules of oxygen

6. What is a capillary? (relevant)
   2 points max
   Small blood vessel
   Blood vessel far from heart
   Site of gas, O2, CO2, transfer
   Small vein/artery
   Where arteries turn to veins or arteries and veins meet

7. How many continuous, closed circuits of blood are there from the heart? (relevant)

1 point

Two (closed) circuits

7b. Name the circuit(s). (relevant)

2 points

Systemic (body)
Pulmonary (lungs)

8. What is an artery? (relevant)

(2) points max (might be in one statement, e.g., carry blood from heart = 2pts b/c each is a point)

Blood vessel
Travels from heart to body
Carries oxygenated blood

9. What is an atrium? (relevant)

(2) points max (might be in one statement, e.g., upper chamber of heart)

Chamber, region, section, part, cavity etc. (of heart)
Blood enters first
Above ventricles, top, upper
Part of heart that blood passes through

10. What is a ventricle? (relevant)

(2) points max (might be in one statement)

Chamber, region, section, part, cavity etc. (of heart)
Blood arrives from atrium
Lower portion of heart
Blood exits heart here

11. What is another name for the left atrioventricular valve?

1 point

Mitral valve
Bicuspid valve

12. Where does the blood entering the left atrium come from? (relevant)

1 point

The lungs
Pulmonary veins
Pulmonary vessels

13. Where does blood entering the left ventricle come from? (relevant)

1 point

Left atrium

14. Where does blood entering the right ventricle come from? (relevant)

1 point

Right atrium

15. The pacemaker is the common term for what specific part of the heart?

1 point

SA node

15b. Where is the "pacemaker" located?

1 point

Located in right atrium

16. Which side of the heart is larger? Why? (relevant)

2 points

Left side of heart is larger
Has larger of 2 circulation routes

17. What is unusual about the pulmonary vein?

1 point

Carries oxygenated blood to heart
Only vein that carries oxygenated blood

18. What are the names of the main veins which carry blood back to the heart from the body?

2 points

superior & inferior vena cava

18b. How many such veins are there?

1 point

two such veins

18c. From what part of the body does each such vein return blood?

2 points

top half (superior) & bottom half (inferior)

HRT_

1. How many valves are found in the heart and where are they located?
   5 points
   Four valves (1 point)
   1 Located between Left atrium & ventricle (1 point)
   1 Located between Right atrium & ventricle(1 point)
   1 Located between Right ventricle & pulmonary artery(1 point)
   1 Located between Left ventricle & aorta(1 point)
   Also accept **2 between atria and ventricle/2 at bottom of atria** (1 point) and **two between ventricles and arteries** (1 point)

2. Where does the blood go after it leaves the left ventricle?
   1 point max
   The body
   Aorta

3. The sound of the heartbeat is often characterized "lub-dub". To what movements does the "lub" correspond? The "dub"?
   4 points max
   LUB = ventricles contracting or emptying (1 point) **and atria closing** (1 point)
   DUB = ventricles relax or fill (1 point) **and valves to arteries snap shut** (1 point)
   If only LUB = heart contract & 1 point only◊DUB = heart expands
   ½ point each if non-specific (must answer which chambers contracting and which valves contracting)

4. What happens when the heart relaxes after a contraction?
   1 point
   Atria fill with blood (heart gets larger)
   Ventricular valves close
   Valves close = ½ point

5. Where does the blood leaving the right ventricle go?
   1 point max
   The lungs
   Pulmonary artery

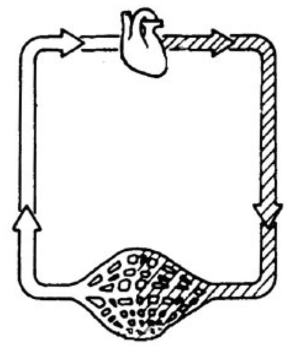
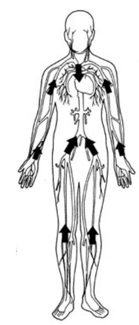
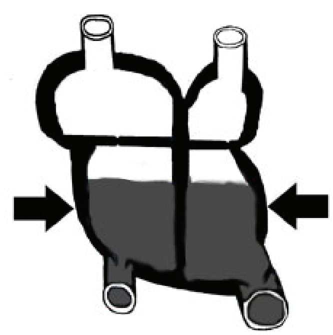
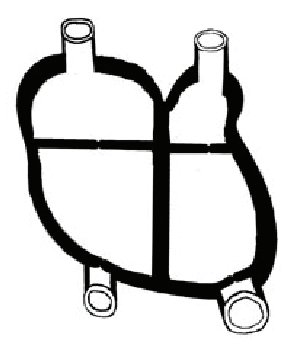
6. What is the solid wall that separates the left and right sides of the heart called?
   1 point
   Septum

7. What color is blood in the left atrium? Why is it that color??
   2 points max
   (Bright) red
   it's oxygenated (fresh), oxygen rich, O2, clean
   If just red = ½ point

8. What color is blood in the right atrium? Why is it that color??
   2 points
   Dark colored (blue, brown, or dark red)
   it's low in O_2, high in CO_2, deoxygenated, dirty, used up

9. Which part of the heart is rounded?
   2 points
   Top (atria) is rounded (.5 if specify right/left top/atria is rounded)
   Bottom (ventricles) is cone shaped (.5 if specify right/left bottom/ventricle cone shaped)

HRP_

Please draw in the circulatory loop including the lungs as it was drawn in the lesson as best you can: *3 points* **Arrows counterclockwise** **Loop comes out of and returns to heart separate from existing loop** **Lungs drawn in connected to second loop (not part of existing loop)**	
What part of the circulatory system and part of the bloods path was this diagram used to illustrate in the lesson? *2 points* **Blood vessels, veins, artery, capillaries** **Return to heart**	
Please indicate what part of the heartbeat (noise) this picture represents. *1 point* **lub** Please label the pools of blood indicating which is red and which is blue. *2 points* **Red on right** **Blue on left**	
Please draw an arrow showing where blood leaves the heart to go to the lungs and where it comes back in *2 points* **Going out (down) of bottom left valve** **Going in (down) into top right valve** *(0 points if more than two arrows -- unless the two critical arrows are labeled correctly)*	

HTT_

What would be the consequences of a large hole in the septum that separates the left and the right ventricles?

3 points max

Mix O2 & CO2 blood
Poor circulation
Poor O2 distribution to body and O2 blood to lungs
Fatigue / lack of energy
When open, blood would not flow through valve without force.

Why doesn't blood pool up in the feet and the lower extremities of the body?

1 point max

Circulatory system is closed or continuous—heart pumping moves blood through system
Valves keep blood flowing in one direction
Muscles w/ veins to push blood up

What would happen if the valves leading out of the ventricles didn't close properly?

2 points max

Backflow into ventricles
Poor overall circulation (to lungs & to body)
Inefficient pumping of blood to body

Imagine that the tempo of contraction of the various parts of the heart was somehow disturbed. What specific implications might this have for the flow of blood through the heart?

2 points max

General disruption of blood flow or rate
Blood accumulation in heart (backup) if fails to contract
Backflow of blood if valves fail to close

What effect, if any, would there be on the efficiency of blood circulation if for some reason the valve between the right atrium and ventricle were unable to close completely?

3 points max

Backflow into right atrium
Poor flow to lungs
Poor overall circulation
Lowered blood pressure

HTP_

What would be the effect on functioning of the circulatory system as a whole if the flow of blood and direction of valves was reversed ?
Text related
1pt max
Everything would be reversed
Little effect, it would still work.
No problems such as death, fatigue, poor circulation etc.

The lesson describes the 'thump' that you hear every time the heart goes against the chest wall. It also describes the 'lub-dub' sound of each beat. Is the thump more likely to coincide with the 'lub' or the 'dub'?
1 pt
Dub (muscle expands, thumping chest wall)

What type of blood vessel is best for transferring gas from red blood cells to parts of the body? What parts of the body exchange gas with blood?
3 pts
Capillaries
Lungs or Alveoli
Extremes (parts of body)

Please describe the part of the heart that is like a funnel. How does that relate to the flow of blood to the body?
2 points
The bottom (ventricle) or left ventricle of the heart is cone shaped
Funnels blood out to body
Gravity assists emptying of heart

Describe in detail the muscle movement in the heart that forces the opening of the valves.
2 points max
Top and bottom heart muscles contract separately
Atrium contracting forces valves between atrium and ventricles open; flaps go out/down
Ventricle contracts, closes those, opens valves to aorta/pulmonary vein
Anything else about contractions = 1 point

APPENDIX B:

Lightning Lesson

Please write down the eight steps of how lightning works.
8 points max
Air rises
Water condenses
Water and crystals fall
Wind is dragged downward
Negative charges fall to the bottom of the cloud
The leaders meet
Negative charges rush down
Positive charges rush up

LRT_

What happens to water vapor as it forms a cloud?
1 point
It condenses into water droplets
Begins to fall
Freezes into crystals

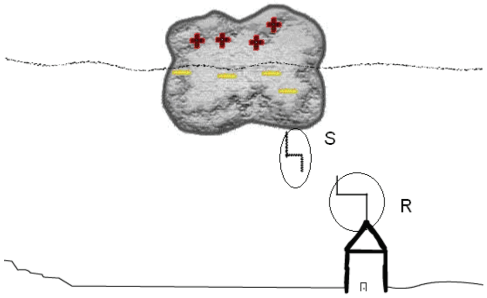
What does only the upper portion of a cloud contain that is important to the formation of lightning?
2 points max
Ice crystals
Positive particles

What is created by rising and falling air current?
2 points max
Electrical charges build
Electrical charges separate
Wind gusts

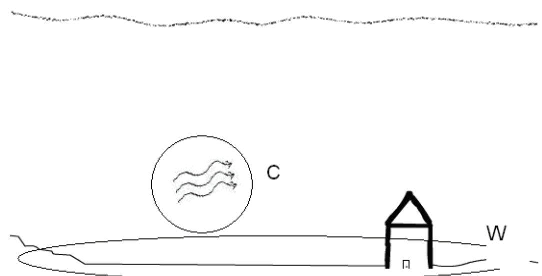
According to the lesson, about how far above the ground do the negative and positive leaders meet?
1 point
165 feet

Does the positive leader come up from the ground or down from the cloud?
1 point
Up from the ground

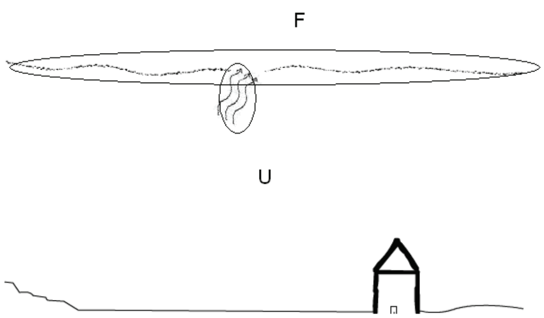
LRP_



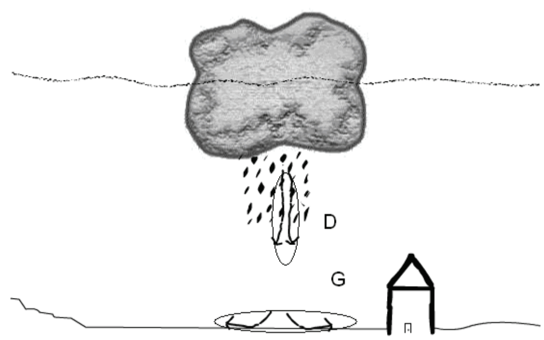
1 Circle the stepped leader and write S next to it. Circle the return stroke and write R next to it.



2 Circle cool moist air and write C next to it. Circle the warmer surface and write W next to it.



3 Circle the updraft and write U next to it. Circle the freezing level and write F next to it.



4 Circle the downdraft and write D next to it. Circle the gusts of cool wind and write G next to it.

LTT_

What could you do to decrease the intensity of lightning?
2 points max
Remove positive particles from earths surface
Placing positive particles near the cloud
Lessen charge differential between ground and sky
1 point for anything related to positive particles

Suppose you see clouds in the sky, but no lightning. Why not?
2 points
Top of cloud not above freezing
No ice crystals form

What does air temperature have to do with lightning?
4 points max (of 6 possible items)
(1) Clouds created by interactions between (2) earths warm surface and (3) oncoming cool air *(2 max from these three)*
(4) Top of cloud above freezing level and (5) bottom of cloud below (6) leads to formation of particles *(2 max from these three)*

LTP_

Movement of which kind of particles are the direct cause of the visible flash of lightning?
1 point
Positive particles

How are the gusts of wind related to the formation of negatively charged particles? Which occurs first?
4 points
Wind causing formation also causes gusts of wind
Wind makes the ice and rain rub and form particles
Wind pushes particles down or separates negative from positive

Negative particles created/formed first

Which leader (negative or positive) occurs first according to the lesson? Which direction (up from earth or down from cloud) does each go?
3 points
Negative first
Negative down
Positive up

APPENDIX C:

NASA-TLX

Please complete the following items to the best of your ability by circling one line on each scale:

MENTAL DEMAND

How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Low **High**

PHYSICAL DEMAND

How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low **High**

TEMPORAL DEMAND

How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was that pace slow and leisurely or rapid and frantic?

Low **High**

PERFORMANCE

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Good **Poor**

EFFORT

How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low **High**

FRUSTRATION LEVEL

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Low **High**

APPENDIX D:

Subjective Questionnaire

If I could change the rate of presentation, I would make the heart lesson:

___Much Faster ___A Little Faster ____It was fine ____ A Little Slower ______Much Slower

If I could change the rate of presentation, I would make the lightning lesson:

___Much Faster ___A Little Faster ____It was fine ____ A Little Slower ______Much Slower

I feel that I was able to understand _____ of the information contained in the heart lesson

___All ____Most ___Some ___Little ____None

I feel that I was able to understand _____ of the information contained in the lightning lesson

___All ____Most ___Some ___Little ____None

I feel that in the heart lesson I was able to determine how the diagram was related to the description

___Strongly Agree ___Somewhat Agree ___Neutral ___Somewhat Disagree ___Strongly Disagree

I feel that in the lightning lesson I was able to determine how the animation was related to the description

___Strongly Agree ___Somewhat Agree ___Neutral ___Somewhat Disagree ___Strongly Disagree

How do you feel you did on the heart exam?

___Very Well ___Well ___OK ___Poorly ___Very Poorly

How do you feel you did on the lightning exam?

___Very Well ___Well ___OK ___Poorly ___Very Poorly

APPENDIX E:

Background Questionnaire

INSTRUCTIONS:

Please complete the following information about yourself and your background.

1. Sex (circle one): M or F

2. Age: _______

3. Approximate GPA: _______

4. Academic year:
   _______ Freshman
   _______ Sophomore
   _______ Junior
   _______ Senior
   _______ Other (please specify) ___________________________________

5. Native Language:
   _______ I'm a native English speaker _______ I'm NOT a native English speaker

6. If not a native speaker, do you consider yourself fluent in English?
   ____Yes ____No

7. Major:
   _______ Psychology _______Other (please specify) _______________

8. Are you Pre-Med?
   ____Yes ____No

9. What Biology courses have you taken in the past (including high school), if any?

Course Number	**Course Name**	**Grade**	**Semester/Year**	**Where?**

10. Have you used multimedia content (e.g., pod casts, instructional CDs, course specific web sites, computer simulations) directly related to course work at Georgia Tech?
    ____Yes ____No

11. Whether you use them or not, please estimate the usefulness of multimedia content in helping you learn.

___Very useful __Quite useful __Neutral __Quite Useless __Totally Useless __No opinion

REFERENCES

Ainsworth, S. E. (1999). A functional taxonomy of multiple representations. Computers and Education, 33(2/3), 131-152.

Ainsworth, S. E., & Fleming, P. (2006). Evaluating authoring tools for teachers as instructional designers. Computers in Human Behavior, 22(1), 131-148.

Ainsworth, S. E., & VanLabeke, N. (2004). Multiple forms of dynamic representation. Learning and Instruction, 14(3), 241.

Anderson, J. R. (1995). Learning and Memory: An integrated approach. New York, NY: John Wiley & Sons, Inc.

Anderson, J. R., & Bower, G. H. (1973). Human associative memory. Washington: Winston and Sons.

Baddeley, A. D. (2003). Working memory: looking back and looking forward. Nature Reviews Neuroscience, 4(10), 829-839.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the Concept of Working Memory. Neuropsychology, 8(4), 485.

Baggett, P. (1989). Understanding visual and verbal messages. In H. Mandl & J. R. Levin (Eds.), Knowledge acquisition from text and pictures. (pp. 101-124): North-Holland.

Barnard, P. J. (1999). Interacting Cognitive Subsystems: Modeling working memory phenomena within a multiprocessor architecture. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control. (pp. 298): Cambridge University Press.

Bieger, G. R., & Glock, M. D. (1986). Comprehending spatial and contextual information in picture-text instructions. Journal of Experimental Education, 54(4), 181.

Bishop, M. J., & Cates, W. M. (2001). Theoretical foundations for sound's use in multimedia instruction to enhance learning. Educational Technology Research & Development, 49(3), 5-22.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. Learning and Instruction, 14(3), 325-341.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. Cognitive Psychology, Vol. 2(4), 331-350.

Brunken, R., Plass, J. L., & Leutner, D. (2004). Assessment of Cognitive Load in Multimedia Learning with Dual-Task Methodology: Auditory Load and Modality Effects. Instructional Science, 32(1), 115-132.

Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: the role of working memory in multimedia learning experiences. Applied Cognitive Psychology, 20(7), 917-940.

Butcher, K. R. (2006). Learning From Text With Diagrams : Promoting Mental Model Development and Inference Generation. Journal of Educational Psychology, 98(1), 182-197.

Byrne, M. D., Catrambone, R., & Stasko, J. T. (1999). Evaluating animations as student aids in learning computer algorithms. Computers & Education 33(4), 253-278.

Chambliss, M. J., & Calfee, R. C. (1998). Textbooks for learning: Nurturing children's minds. Malden, Mass.: Blackwell Publishers

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition & Instruction, 8(4), 293.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. British Journal of Educational Psychology, 62(2), 233.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149-210.

Cook, L. K., & Mayer, R. E. (1988). Teaching Readers about the Structure of Scientific Text. Journal of Educational Psychology, 80(4), 448-456.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. Cognitive Psychology, 6(1), 84-107.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychological Bulletin, 104(2), 163.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning & Verbal Behavior, Vol. 11(6), 671-684.

De Beni, R., Pazzaglia, F., Gyselinck, V., & Meneghetti, C. (2005). Visuospatial working memory and mental representation of spatial descriptions. European Journal of Cognitive Psychology, 17(1), 77-95.

Dubois, M., & Vial, I. (2000). Multimedia design: the effects of relating multimodal information. Journal of Computer Assisted Learning, 16(2), 157.

Eskicioglu, A. M., & Kopec, D. (2003). The Ideal Multimedia-Enabled Classroom: Perspectives from Psychology, Education, and Information Science. Journal of Educational Multimedia & Hypermedia, 12(2), 199.

Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10(2), 181-197.

Gagne, R. (1972). Domains of learning. Interchange 3(1), 1-8.

Gagne, R., Briggs, L., & Wager, W. (2005). Principles of Instructional Design (5th ed.). Fort Worth, TX: HBJ College Publishers.

Gellevij, M., Van Der Meij, H., De Jong, T., & Pieters, J. (2002). Multimodal versus unimodal instruction in a complex learning context. Journal of Experimental Education, 70(3), 215.

Ginns, P. (2005). Meta-analysis of the modality effect. Learning & Instruction, 15(4), 313-331.

Good, T. L., & Brophy, J. E. (1990). Educational psychology: A realistic approach (4th ed.): Longman/Addison Wesley Longman, Inc.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning About Relations. Psychological Review, 112(2), 468–493.

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.

Harskamp, E. G., Mayer, R. E., & Suhre, C. (2007). Does the Modality Principle for Multimedia Learning Apply to Science Classrooms? Learning and Instruction, 17 (5), 465-477.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. Learning and Instruction, 343-351.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. Journal of Educational Psychology, 84(1), 76-84.

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. Journal of Educational Psychology, 87(1), 18-32.

Jeung, H. J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. Educational Psychology, 17, 329-343.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. Applied Cognitive Psychology, 13(4), 351.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control. (pp. 183): Cambridge University Press.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. Journal of Research in Science Teaching, 34(9), 949-968.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11(1), 65.

Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. Applied Cognitive Psychology, 17(4), 401.

Low, R., & Sweller, J. (2005). The modality principle. In R. Mayer (Ed.), Cambridge Handbook of Multimedia Learning (pp. 147-158). New York: Cambridge.

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. (1989). Systematic Thinking Fostered by Illustrations in Scientific Text. Journal of Educational Psychology, 81(2), 240-246.

Mayer, R. E. (2001). Multimedia Learning. Boston: Cambridge University Press.

Mayer, R. E. (2005). Cognitve Theory of Multimedia Learning. In R. E. Mayer (Ed.), The Cambridge Handbook of Multiemdia Learning (pp. 30-48). New York: Cambridge University Press.

Mayer, R. E. (2005). Principles for managing essential processing in multimedia learning: segmenting, pre-training, and modality principles In R. E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning (pp. 169 - 182). New York: Cambridge.

Mayer, R. E., & Anderson, R. B. (1991). Animations Need Narrations: An Experimental Test of a Dual-Coding Hypothesis. Journal of Educational Psychology, 83(4), 484-490.

Mayer, R. E., & Anderson, R. B. (1992). The Instructive Animation: Helping Students Build Connections between Words and Pictures in Multimedia Learning. Journal of Educational Psychology, 84(4), 444-452.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, 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.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? Journal of Educational Psychology, 93(2), 390-397.

Mayer, R. E., & Gallini, J. K. (1990). When Is an Illustration Worth Ten Thousand Words? Journal of Educational Psychology, 82(4), 715-726.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. Journal of Educational Psychology, 93(1), 187-198.

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. (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., Boire, M., & Vagge, S. (1999). Maximizing Constructivist Learning From Multimedia Communications by Minimizing Cognitive Load. Journal of Educational Psychology, 91(4), 638-643.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. Journal of Educational Psychology, 86(3), 389.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. Educational Technology Research and Development, 43(1), 31-43.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63(2), 81-97.

Miyake, A., & Shah, P. (1999). Models of working memory : mechanisms of active maintenance and executive control. Cambridge ; New York: Cambridge University Press.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. Journal of Educational Psychology, 91(2), 358-368.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. Journal of Educational Psychology, 92(1), 117-125.

Mousavi, S. Y., Lowe, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87(2), 319.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. Educational Psychologist, 38(1), 63-71.

Paivio, A. (1986). Mental representations: A dual coding approach. New York: Oxford University Press.

Palmiter, S. L., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. Human-Computer Interaction, 8(3), 193-216.

Penney, C. G. (1975). Modality effects in short-term verbal memory. Psychological Bulletin, 82(1), 68.

Penney, C. G. (1980). Order of report in bisensory verbal short-term memory. Canadian Journal of Psychology, 34(2), 190.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. Memory & Cognition, 17(4), 398.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. Journal of Experimental Psychology: Applied, 7(3), 219-226.

Salomon, G. (1994). Interaction of media, cognition, and learning. Hillsdale, N.J.: L. Erlbaum Associates.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? International Journal of Human-Computer Studies, 45(2), 185-213.

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

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. Learning & Instruction, 13(2), 141.

Schnotz, W., Bannert, M., & Seufert, T. (2002). Toward an integrative view of text and picture comprehension: Visualization effects on the construction of mental models. In Psychology of science text comprehension. (pp. 385): Lawrence Erlbaum Associates, Publishers.

Stasko, J., Catrambone, R., Guzdial, M., & McDonald, K. (2000). An evaluation of space-filling information visualizations for depicting hierarchical structures. International Journal of Human-Computer Studies, 53(5), 663-694.

Stenning, K. (1998). Representation and conceptualisation in educational communication. In M. W. van Someren, P. Reimann, H. P. A. Boshuizen & T. d. Jong (Eds.), Learning with multiple representations (pp. 320-333). Amsterdam: Pergamon.

Sweller, J. (1999). Instructional Design. Melbourne: ACER Press.

Sweller, J. (2005). Implications of Cognitve Load Theory for Multimedia Learning. In R. E. Mayer (Ed.), The Campbridge Handbook of Multimedia Learning (pp. 19-30). New York: Cambridge University Press.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive Load as a Factor in the Structuring of Technical Material. Journal of Experimental Psychology: General, 119(2), 176-192.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10(3), 251.

Tabbers, H. K. (2002). The modality of text in multimedia instructions. Open University of the Netherlands, Herleen, Netherlands.

Tabbers, H. K., Martens, R. L., & van Merrienboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. British Journal of Educational Psychology, 74(1), 71-81.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. Behavior Research Methods, 37, 498-505.

Weinstein, L. F. (1987). Instruction for military air intercept control. Paper presented at the 31st Annual Meeting of the Human Factors Society, Santa Monica, CA.

Wickens, C. D., & Hollands, J. (2000). Engineering psychology and human performance (Vol. 3). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. Human Factors, 30(5), 599-616.

Wolfe, M. B. W., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W., et al. (1998). Learning from text: Matching readers and texts by latent semantic analysis. Discourse Processes, 25(2), 309-336.

Zimmer, H. D. (2004). The Construction of Mental Maps Based on a Fragmentary View of Physical Maps. Journal of Educational Psychology, 96(3), 603-610.

Zolna, J. S., & Catrambone, R. (2007). The Effects of Study Time and Presentation Modality on Learning. Paper presented at the Human Factors and Ergonomics Society 51st Annual Meeting, Baltimore, MD.