ABSTRACT

Title of Document: AN EXAMINATION OF THE IMPACT OF
                                            COMPUTER-BASED ANIMATIONS AND
                                            VISUALIZATION SEQUENCE ON
                                            LEARNERS' UNDERSTANDING OF
                                            HADLEY CELLS IN ATMOSPHERIC
                                            CIRCULATION

                                            Daniel Wyatt Harris, Ph.D., 2012

Directed By: Professor William G. Holliday, Department of
                                            Teaching and Learning, Policy, and Leadership

Research examining animation use for student learning has been conducted in the

last two decades across a multitude of instructional environments and content areas. The

extensive construction and implementation of animations in learning resulted from the

availability of powerful computing systems and the perceived advantages the novel

medium offered to deliver dynamic representations of complex systems beyond the

human perceptual scale. Animations replaced or supplemented text and static diagrams of

system functioning and were predicted to significantly improve learners' conceptual

understanding of target systems. However, subsequent research has not consistently

discovered affordances to understanding, and in some cases, has actually shown that

animation use is detrimental to system understanding especially for content area novices (Lowe 2004; Mayer et al. 2005).

This study sought to determine whether animation inclusion in an authentic learning context improved student understanding for an introductory earth science concept, Hadley Cell circulation. In addition, the study sought to determine whether the timing of animation examination improved conceptual understanding. A quasi-experimental pretest posttest design administered in an undergraduate science lecture and laboratory course compared four different learning conditions: text and static diagrams with no animation use, animation use prior to the examination of text and static diagrams, animation use following the examination of text and static diagrams, and animation use during the examination of text and static diagrams. Additionally, procedural data for a sample of three students in each condition were recorded and analyzed through the lens of self regulated learning (SRL) behaviors. The aim was to determine whether qualitative differences existed between cognitive processes employed. Results indicated that animation use did not improve understanding across all conditions. However learners able to employ animations while reading and examining the static diagrams and to a lesser extent, after reading the system description, showed evidence of higher levels of system understanding on posttest assessments. Procedural data found few differences between groups with one exception---learners given access to animations during the learning episode chose to examine and coordinate the representations more frequently. These results indicated a new finding from the use of animation, a sequence effect to improve understanding of Hadley Cells in atmospheric circulation.

AN EXAMINATION OF THE IMPACT OF COMPUTER-BASED ANIMATIONS
AND VISUALIZATION SEQUENCE ON STUDENT UNDERSTANDING OF
HADLEY CELLS IN ATMOSPHERIC CIRCULATION

By

Daniel Wyatt Harris

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:

Professor William G. Holliday, Chair
Professor Robert F. Marcus
Professor J. Randy McGinnis
Professor Denis F. Sullivan
Professor Olivia N. Saracho

Acknowledgements

In the course of completing this research project, many individuals provided invaluable guidance and support both intellectually and morally over the years. First, I want to recognize my advisor, Dr. William G. Holliday, for his patience and wisdom. This opportunity and project would not have come to fruition without his advocacy, insight and ability to guide me through the process of refining a research project and seeing it through to completion given the inevitable barriers encountered in any long term endeavor. I would also like to thank Dr. Roger Azevedo for introducing me to the science of human cognition and enabling me to identify a research topic of personal interest and application in my own teaching career. I am indebted to Dr. Azevedo for providing the guiding framework through which I am able to offer this small contribution. I would also like to thank Dr. J. Randy McGinnis for encouraging me to discover my own writing position and for providing specific and important recommendations for the improvement of this document. I would also like to recognize Dr. Robert Marcus, Dr. Denis Sullivan and Dr. Olivia Saracho for their willingness to serve on my doctoral committee and the insightful contributions they made through their careful reading of this document. These scholars offer an impressive model to emulate through the remainder of my career.

For moral and logistical support, I would like to thank my family, friends and colleagues. All have contributed in more ways than I can recall in my ability to commute to College Park over the years and to complete this project. I would especially like to recognize my wife, Jill, and our children, Ayden and Solomon. I would have never completed this work without her support, sacrifice and prodding. She pushed me through

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Given the proliferation of computer projection systems in higher education science classrooms, and the technological innovations enabling textbook publishers to develop animations representing complex physical systems inexpensively, most introductory science texts include an extensive array of animated sequences available for instructional delivery and learner consultation. Animation sequences included in textbook packages are constructed to provide dynamic audiovisual representations of content described in the text and with reference to static diagrams to aid student understanding and facilitate the conceptual understanding of dynamic processes within systems. Most animations are designed to externally represent complex interactions in dynamic physical systems due to the explicit and implicit advantages assumed of the "active" animation medium (Tversky, Morrison, & Bétrancourt, 2002). These advantages include translations and transformations of system components in space and time, cause and effect propagations through system components, and the representation of components that can only be viewed at microscopic and macroscopic scales.

While the inclusion of animation in instruction can provide a "wow factor" in the classroom and an attractive feature for online learning, there are still mixed results as to the effect of animation on student conceptual understanding. Studies find that animation use increases understanding in certain content areas (Höffler & Leutner 2007) including particulate models in chemistry (Russell, Kozma, Jones, Wykoff, Marx & Davis 1997), additional studies including a meta-analysis find that animation inclusion in itself does not improve conceptual understanding (Tversky et al. 2002). Similar to many classroom

innovations, the absence of increased student understanding is most often attributed to ineffective animation design, inappropriate deployment in the instructional setting and the discovery that animations often contain more content information than the text/static diagrams (Mayer, Hegarty, Mayer & Campbell 2005; Narayanan & Hegarty 2002; Lowe & Schnotz 2005). Given that most prior research fails to examine how animations are employed during the delivery of instruction in undergraduate (with the notable exception of Velazquez-Marcano et al. 2004) , introductory science content (Baek & Layne 1988; Park & Gittelman 1992; Williamson & Abraham 1995), this research focuses on whether the timing of the presentation of animations can be associated with increased comprehension.

To date, most animation learning research has taken place in controlled laboratory environments (primarily educational psychology labs) using undergraduate education and psychology majors. These settings do not mimic classroom environments and typical learning settings as they often require subjects to report on experiences within a cubicle outside of the more contextually common learning environment on concepts that are disconnected from a specific course or discipline. This lack of ecological validity, or authenticity, and reduction in likelihood of a true representative sample of undergraduate learners can be hypothesized to reduce the generality with which some prior research findings can be interpreted given subject characteristics (Brewer 2000). While these studies provide important information about animation design that can facilitate learning and illuminate how selected undergraduates learn in mostly brief, personal computer-human interactions, these settings differ dramatically from the context of an actual undergraduate science classroom and laboratory. These shortcomings have been

highlighted in the research examining animation used in introductory earth science concepts, resulting in a shift to treatments in authentic classrooms (Maher 2002; Lowe 2004; Stull & Maher 2007; Yezierski & Birk 2006).

This research study sought to extend prior animation research by examining two important considerations related to their use in an undergraduate, introductory earth science lecture and laboratory course. First, the study examines whether the use of textbook publisher animations depicting atmospheric circulation processes improves content understanding in a classroom instruction context. While Lowe (2003) and Edelson and Gordin (1998) have examined animation and computer delivered instruction in the earth science context, researchers have not evaluated textbook produced animations even though these resources are included in the typical undergraduate textbook package or are available on textbook support web sites. Second, this study examines pedagogical implications of animation use by assessing whether the timing of animation use by learners impacts conceptual understanding. Given the ease with which animations can be incorporated into instructional and learning sequences in modern classrooms and laboratories, the instructor and/or learner must decide when to display or view animations to facilitate the coordination of all available representations of the target concept (e.g., text and static diagrams). To date, most animation research overlooks the question of sequencing and timing during instruction and learning instead focusing upon animation design within multimedia environments to maximize learner understanding (Mayer 2001) and any contrasting content between different representation types (e.g., static diagrams versus animations) (Mayer et al. 2004).

The first research question asks: Does the inclusion of animations during the learning episode result in increased student conceptual understanding of tropical atmospheric circulation and if so, the extent of this gain? The second research question asks: If animations are found to be beneficial to student learning in the content area, does the timing of the presentation of the animation sequences affect student understanding and comprehension of the examined system? Three temporal states for animation delivery were defined for this research study including 1) an introduction to the target content prior to a textual, static or didactic examination of the components and functions of the system, 2) a summarization viewed following the textual and static explanations, or 3) user-selected timing in which the animation sequences may be embedded within the overall presentation of the learning materials and activated as needed by the learner when seeking to comprehend the component properties and functions of the circulation system.

The research design of this study dictated three specific experimental conditions related to animation use, contrasted against a control group in which an animation sequence is not viewed during the learning episode. Ecological validity issues in the research design were mitigated by implementing the pedagogical experiment during a component of an introductory, undergraduate lab science course. While the sample was dictated by student registration for the course rather than random assignment from the entire student population, subjects were more heterogeneous than most prior research by expanding the potential sample representatives beyond introductory education and psychology students and/or paid or recruited participants.

A secondary goal of this research project was to compare and contrast the cognitive processes enacted by learners in each treatment condition to examine how the timing of the animation sequences affected their ability to understand the target concepts. A small random, sample of learners (n=3) from each treatment condition and the control group was selected to produce "think alouds" during the learning episode to illuminate the cognitive strategies employed to coordinate the three representational content types available. Differences between each condition and individual are discussed in an effort to evaluate optimal timing for learner comprehension in support of the statistical findings.

Chapter Two: Literature Review

Section 1. Learning Theory, Animation and Information Processing Theory

The goal of this study is to determine whether animations aid understandings of tropical atmospheric circulation (i.e., Hadley Cells) and whether the timing of animations viewership impacts the extent of conceptual understanding or learning. Learning in the context of this study refers to classroom and laboratory knowledge acquisition and construction derived from prior knowledge, didactic instruction, textbook reading, diagram examination and animation review. Learning from content portrayed in animations has actually been the focus of continuing research for over two decades. Most researchers in this domain design and interpret their studies through a constructivist theoretical and methodological framework of Information Processing Theory (IPT) which seeks to understand the cognitive processes employed during learning (Mayer 1996; Hegarty 1992; Azevedo 2004). Animation researchers utilizing the IPT model of learning seek to develop and evaluate content representations and features in animations to align these features to human sensory inputs and cognitive processes to facilitate learning (Miller 2003).

The IPT model straddles two of the three leading philosophical perspectives of learning in the educational psychology community: cognitivism and constructivism (Mayer 1996). The first of these three leading theories, behaviorism, is based on the supposition that learning occurs with a change in behavior, is shaped by the environment, and is guided by the principles of contiguity and reinforcement (Skinner 1950). Cognitivism supersedes the first model adapted by research psychologists, behaviorism, by refuting the conceptualization of learning as solely a product of stimulus-response

associations and extends the conceptualization of learning to encompass not just behavior modification based on rewards and punishments but the realization that cognitive processing of external information leads to understanding and productive thinking, not just reproduction, enabling transfer to novel, problem-solving situations (Mayer 1996). Constructivists' extended the cognitivism view and defined learning as an individual construct resulting from personal interpretations of all life experiences including both formal and informal education. Through time, constructivist tenets expanded to include the importance of social guidance and interactions to make meaning of life and educational experiences.

This chapter begins with a review of learning theory as applied to education and learning, with a detailed examination of the three leading learning theories and reference to the models that have followed. The chapter continues by reviewing learning analysis performed in science education in general and the IPT model in particular and explains how the IPT approach links to the animation research and analysis used in this study. The chapter concludes by placing this study within the context of the literature that has informed this study.

Section 2. Three Leading Perspectives on Learning

Behaviorism is often cited as the dominant learning theory for the first half of the 20th century (Doolittle & Camp 1999; Thorndike 1932; Wirth 1972). In this model, learning was understood as the crystallization between stimuli and responses in which rewards were provided to reinforce the behavior (Camp 1983). The classic example of learning as response strengthening in the behaviorism model was Pavlov's experiment

using bells to initiate dog salivation in place of a food dish once conditioning was achieved (Dembo 1994). Behaviorists contended that human actions (or behavior) could be explained by observable phenomena in their environment causing specific reactions thereby solidifying responses into reflexive action (Camp 1983). Proponents of behaviorism designed educational environments to teach learners through work and moral habits, including well defined drill and practice lessons with the expected outcome of voluntary adoption of the behaviors explicitly rewarded in classrooms (Mayer 1996). Through time and observational research, a competing learning model was suggested after researchers recognized and acknowledged that different types of learning also existed. The new model, cognitivism, extended the conceptualization of learning to encompass not just behavior modification based on rewards and punishments but the realization that cognitive processing of external information leads to understanding and productive thinking, not just reproduction, enabling transfer to novel, problem-solving situations (Mayer 1996).

Cognitivist researchers understood learning to result from information acquired from external inputs and the cognitive processing of the information delivered by these inputs with stored knowledge of related information in memory to reorganize the resultant information in new cognitive structures (or schema) (Good & Brophy 1990). Researchers employing this new framework devoted most efforts to lab-based research and simple learning tasks disconnecting the subject from the confounding influences of real world scenarios. Cognitivism's lab-based paradigm was criticized for a lack of ecological validity and the omission of research on learning in authentic academic settings. Psychologists studying learning reacted to this criticism by shifting their focus to

education research and began to examine learning specific to academic tasks and disciplines resulting in a modified form of cognitivism termed constructivism.

Constructivism argued that academic learning was a complex and varied process of knowledge construction different than simple lab-based problem-solving (Mayer 1996). Moreover, learners' context and individual differences were understood to impact knowledge acquisition and construction therefore resulting in less objective and more open-ended models of instruction and learning (Merrill 1991). Constructivists view learning as an individual construct resulting from personal interpretations of all life experiences including formal and informal education thus presenting the learner as sense-maker (Steffe & Gale 1995). As research proliferated in authentic learning environments, constructivist tenets expanded based on epistemological beliefs concerning the nature of knowledge resulting in the formulation of continuum of theoretical viewpoints through which learning can be interpreted (Doolittle & Camp 1999). At one end of the continuum, cognitive constructivism views knowledge as an external, independent reality that is knowable by an individual learner and seeks to understand the cognitive processes at work to internalize this knowledge. In contrast, the opposing end of the continuum, radical constructivism, views knowledge as an internal construct resulting from the accumulation of interactions and experiences when navigating the world through life. Thus the internal representation is a model, situated in an individuals' own context, seeking to discern meaning (Doolittle & Camp 1999). Lying between these epistemologies, social constructivism recognizes the importance of social guidance and interactions to make meaning of life and educational experiences. Researchers adopting this framework shifted the focus of learning from processes of the mind or individual

constructs, to the social and cultural environment in which the individual participates (Vygotsky 1978). Experienced society members were seen to convey the important elements of their culture to the new members, the learners, through interactions in both formal and informal ways and both inside and outside of the home, school and community.

Currently, most researchers stake a position along the constructivist continuum and recent learning theories seek to explore the nature of knowledge from one of these three perspectives. Researchers interested in technology and education, including animation researchers, adopted elements from each of these theoretical shifts to design computer-based learning environments while selecting the appropriate theoretical framework based on their specific research aims. Because most instructional animations are delivered via computers in classroom and laboratory settings to the individual learner, most researchers maintain an objective cognitive constructionist framework focusing on learning processes and outcomes related to objective target content. Thus IPT continues to form the theoretical lens through which a great deal of animation research similar to this study is conducted with elaborations specific to multimedia learning evident (Mayer 2001; Hegarty 1992).

The principles comprising Mayer's (2001) *Cognitive Theory of Multimedia Learning* proceed from this theoretical lens and have been empirically evaluated and refined through an extensive, long-term and large-scale research program across academic disciplines as well as learning contexts. Mayer's theory serves as the guide through which the differences observed in this study between the experimental groups are explained. This study attempts to address limitations noted in Mayer's learning

assessments with the collection of procedural or process data for a small sample of learners engaged with a learning activity including multimedia representations in the form of animation. Critics note that Mayer's work infers cognitive processing by the learner without querying or observing the learner directly concerning these processes. These researchers advocate the collection and analysis of process or procedural data, often in the form of think aloud protocols, to reveal inferred learning strategies in multimedia environments. Azevedo's (2004) self-regulated learning research and analytic methodology was chosen to interpret the process data collected in this study in an effort to link the theory of multimedia learning to individual learners' cognitive processes viewed through the lens of self regulated learning behaviors. Based in the IPT framework, this coding method seeks to compare cognitive processes employed by a sample of the study's participants to determine whether differences exist between treatments conditions and if successful strategies for conceptual understanding can be identified.

Section 3. Learning Research in Science Education

Animation research progressed from foundational cognitivist studies examining the mental processes involved in coordinating multi-modal (i.e., audio and visual) sensory inputs to multimedia design recommendations to facilitate conceptual understanding from target content (Baddeley 1992; Hegarty 2003; Mayer 2001). These ongoing lines of research which evaluate learning processes and outcomes in the instructional use of animations have been couched within the instructional practice of teaching for conceptual change, frequently understood as mental model representations

and their changing nature during the learning process (Greene & Azevedo 2007; Human-Vogel 2006; Ke et al. 2005). Researchers adopted mental model assessments to illuminate holistic conceptual understandings of target content before and after treatments rather than more common assessments that simply evaluate change in declarative knowledge.

Science curriculum and instruction researchers utilized conceptual change assessments in the last decade before shifting focus toward the recent framework of learning progressions. In this framework, science curriculum and pedagogy are aligned across grade levels to allow learners to revisit key scientific concepts and principles throughout their formal education, extending and refining scientific knowledge across the separate disciplines with ever increasing sophistication thereby promoting proficiency prior to secondary school (Duschl, Schweingruber & Shouse 2007). However, based on the author's teaching experience with undergraduate non-science majors and current curriculum in K-12 education in Maryland, most students possess little prior knowledge related to the underlying physical science specific to meteorological and climatological processes. Thermodynamic concepts within the earth-atmosphere system underlie the atmospheric circulation content examined in this study and challenge novice learners given the integrative nature of the components of the system.

Section.4. Information Processing Theory

The information processing theory (IPT) of cognitive development seeks to identify and understand the manner in which the human mind receives sensory inputs (or information) from an external source (i.e. the environment) and the cognitive processes

involved in storing, retrieving and performing operations on this information in the domain of memory (Siegler & Alibali 2005). Cognitive processes are analyzed and evaluated within the framework of IPT to understand how the human mind solves problems related to complex, simple and day to day activities or tasks.

Information processing theory states that all environmental input is information and that thinking is the sequential process of evaluating, processing and using the information to comprehend the surrounding world and/or to answer questions developed within the environment (Siegler & Alibali 2005). The interface responsible for "making sense" or converting information into knowledge in a highly complicated environment is the construct of memory where memory is sub-divided into three types: 1) sensory, 2) working (or short-term) and 3) long-term.

Sensory memory is described as the receptor of external stimuli and functions to convert these stimuli into electrical or chemical impulses through which our body's senses can transfer information to the brain (or working memory) for processing. Sensory memory temporarily stores, for between one-half to three seconds, vast quantities of external stimuli, both auditory and visual given adequate attention. Limited quantities of information or chunks, which are defined, discrete aggregations of sensory information, are transferred (or encoded) in working memory due to storage limitations. Research indicates that the number of "chunks" that can be transferred to working memory are approximately seven plus or minus two (Kehoe 1999). Moreover, if any chunk of information is to be retained in short-term memory for more than 18-20 seconds, repetition or rehearsal must performed to ensure that the information is not lost. Alternatively, long-term memory (LTM) is theorized to be unlimited in both storage

quantity and duration although the successful retrieval of knowledge from within LTM is reliant upon the veracity of the network of connections constructed by the individual. Within LTM, four types of knowledge are created by working memory and have been described as declarative, procedural, episodic, and conditional knowledge as related to the types of information individuals are able to successfully retrieve.

Processing and the integration of both sensory inputs and LTM knowledge is undertaken in working memory through conscious thought under the auspices of an executive control. The executive control consumes operational space within working memory, thereby decreasing the quantity of manageable chunks and functions, as a metacognitive tool for monitoring progress towards an individually conceived goal or solution. However, working memory has been shown to streamline repetitive processes, or automatize processes with experience, thereby freeing working memory, maximizing efficiency, and allowing for additional external inputs or information to be processed. Additionally, information transferred to working memory has been shown to fall into two broad categories: auditory inputs and visuospatial inputs. Operations are performed on each type of input simultaneously (working memory can be seen as a parallel processor using a computer metaphor) but separately, an important consideration when describing the perceived advantages of learning with multimedia, specifically animations. Chunk and operational limitations in working memory dictate that complex problems often strain cognitive resources thereby necessitating mental adaptation and/or flexibility when faced with these situations. Problem solving research, often adapted to assess learner comprehension and understanding, seeks to describe methods used to attain goals or seek solutions in a complex world given the limitations of our cognitive systems. The problem

solving space and cognitive processes enacted during problem solving is considered the primary method to generate conceptual understanding (or learning) and has been examined frequently in an effort to illuminate learning from diagrammatic, multimedia and computer-based environments.

The IPT problem-solving framework is applied within the context of animation study because the cognitive processes employed by the learner working within representational environments, including animations, is important to inform content selection, representational style and presentation mode to facilitate conceptual understanding. Thus research conducted by Hegarty et al. (2003), Mayer (2001), Lowe (2008), Azevedo & Cromley (2004) and Edelson & Gordin (1998), integrates cognitive studies with instructional design and delivery to examine the construction of understanding often in the form of mental models when seeking to solve problems related to system representations as portrayed by multimedia representations including animations. Mental model theory, constructs and assessment, are principle components for understanding learning through conceptual change.

Section 5. Mental Animations

Early animation study sought to discern the cognitive processes employed by learners to understand simple mechanical representations or problems based on elementary physics knowledge and/or real world experience with mechanical devices (Hegarty 1992). Once baseline understandings were refined (Hegarty 2000), increasingly complex mechanical systems (e.g., three-dimensions versus two-dimension systems) and computer-based environments were incorporated and served as a basis from which to

develop an explanatory model of individual's ability to perform mental animation. The construction of mental animations is seen as requisite cognitive activity for understanding the spatial and temporal dynamics associated with functioning of complex physical systems. Thus, external representations of these systems in the form of animations should be designed to mimic learner's mental animation processing to facilitate understanding.

The theory of mental animation and subsequent research on external-internal animation cognition was developed and refined over two decades of empirical experimentation on mechanical reasoning (Hegarty & Cate 2003). Hegarty, the leading researcher in this research area, and her colleagues, initially developed static, two-dimensional representations of mechanical systems (e.g., pulley systems, levers, and gears) and posed both static and kinematic questions to undergraduate university students about movements in the system (Hegarty 1992, Hegarty & Steinoff 1997, Hegarty & Kozhevnikov 1999). Through time, Hegarty and colleagues increased system complexity and examined hypermedia representations (including animations) of mechanical systems to elaborate her emerging theory of mental animation (Hegarty, Kriz & Cate 2003; Hegarty, Narayanan & Freitas 2001). The aim of the research was to determine the specific methods by which people encode a static representation of a mechanical system and then mentally animate the system in working memory. By mentally animating the components of a simple, mechanical system, the individual is able to infer the movement of all system components from the movement of just one component given an understanding of motion or kinematics. To successfully perform Hegarty's mental animation tasks, an individual was required to encode relevant information from the external representations (i.e., text and diagram), retrieve prior knowledge and/or internal

representations or mental models of an analogous situation or mechanical device, and apply any algorithms (if familiar with physics) and/or heuristics stored in long term memory to identify a probable solution to the problem (or answer to the question provided). Because even simple mechanical systems included multiple components, several inferences (or component transformations) were required before stating a probable or correct solution.

Initial results from eye tracking studies indicated that participants animated the system in a piecemeal sequence based on the movement of clustered fixation points (or gazes) and because response times increased as distances down the causal chain increased (Hegarty 1992). Error rates were also shown to increase with distance from the stated motion or referent within the system representation further supporting the piecemeal structure of the participant's mental representations. Hegarty integrated these finding into the IPT and problem solving framework to develop a production system model describing how the animation process appears to be undertaken. Hegarty interpreted the piecemeal finding to advance the idea that participants decomposed or created sub-goals within the problem space due to limitations in working memory. Moreover, the gaze data illuminated the notion that participants were only able to transform components or envision motion in components that were spatially contiguous due to their limited knowledge of mechanical systems.

The identification of the significance of prior knowledge about systems to facilitate understanding is important in the context of this study given the differences in the timing of animation viewership between the experimental groups. In the first treatment group, the animations are displayed prior to reading the text and examining the

static diagrams. Thus, with limited or no prior knowledge, the learner is theorized to be less likely to identify the relevant features and interrelationships to encode thereby resulting in little system understanding. Moreover, misconceptions may be generated by viewing the representations in this temporal sequence that may inhibit understanding when presented with verbal and static representations of the phenomena.

Because significant variations between individuals' response times and accuracy measures were noted, the importance of spatial abilities led to work comparing animation understanding between learners with high and low spatial abilities (Hegarty & Steinoff 1997). Spatial abilities, including mental transformation, rotation and perspective taking are considered an important cognitive component to both internal mental animation processes and encoding processes in external animation comprehension. Hegarty and Kozhevnikov (1997) developed predictive models for mental animation performance using spatial abilities testing. Research outcomes supported the importance of intrinsic spatial ability in mechanical reasoning and informed this study by the inclusion of a mental rotation pre-test. Study participants' performance on this pre-test are used to categorize and analyze understanding differences between high and low mental rotation ability subjects in the results and analysis section.

Technological advances and the proliferation of powerful personal computers enabled the presentation of complex mechanical systems to increase dramatically; however, initial studies did not find the expected dramatic gains from animation use (Hegarty and Narayanan 1998). In response, studies shifted focus to evaluate animation design in an effort to determine why the expected learning gains were not occurring and how to alter animation design to produce learning gains.

Section 6. Multiple Representations, Cognitive Load and Design Decisions in Animation

Foundational research evaluating multimedia design and learner processing adapted by animation researchers was spearheaded by Mayer and colleagues' extensive research (2001) informed by cognitive sciences' IPT model of sensory input and working memory. This research focused on the coordination of multiple representations of target information and led to principles of design in multimedia presentations based upon the understanding that human cognition coordinates more than one sensory input simultaneously. Mayer's (2001) work was informed by research illuminating the disctrete channels of sensory inputs. Multiple representations of information for problem-solving or task completion are seen to be superior to single sensory inputs based on Baddeley's (1992) and Clark and Paivio's (1991) work detailing independent and separate auditory and visual processing channels in working memory. Results from their research indicate that the maximum capacity of information (i.e., chunks) in working memory can be increased when sensory inputs are presented by concurrent auditory and visual means.

Contemporaneously, Chandler and Sweller (1991) introduced a theory of limited short-term memory defined as cognitive load in the context of learning. Sweller (1994) later extended the cognitive load theory and identified three sources of short-term memory load: intrinsic cognitive load, extraneous cognitive load and germane cognitive load. Sweller's model states that in a given learning situation or episode, students' short-term memory is taxed by the inherent difficulty of the subject matter or *intrinsic cognitive load* (subject to individual differences due to prior knowledge and experience with the subject matter), the method in which the information is represented (e.g., text, diagrams,

animations, etc.) or *extraneous cognitive load*, and the cognitive functions required to process the information and build mental representations or understanding or *germane cognitive load*. Therefore, the only real load that can be impacted and minimized by curricular designers and instructional specialists (notwithstanding strategy instruction to improve germane load processing) is the presentation or representation of the subject matter content for the learner. Given technological advances enabling instructors to seamlessly present curriculum content by auditory as well as graphical means, design issues related to multimedia or multiple representations spurred extensive research into methods to minimize extraneous loads on the learner (Bodemer 2004; Goldman 2003; Lowe 2004; Mayer 2001; Moreno 1999, 2002, & 2004).

Mayer et al.'s (2001) resultant multimedia model relies upon the empirically tested assumption of two independent and simultaneous input channels, auditory and visual (i.e. Pavio's Dual Coding Theory) which thereby increase the information type and quantity transferred to working memory. Mayer's model, assumed that by using *careful* multimedia design, learners encode both audio and visual (i.e., multiple) representations of the target phenomena in working memory and after retrieving prior (or existing) knowledge from long term memory (LTM), integrate novel information with existing information into an organized cognitive structure termed a mental model in LTM. Long-term memory (LTM) is theorized to be unlimited in both storage quantity and duration although the successful retrieval of knowledge from within LTM is reliant upon the veracity of the network of connections constructed by the individual. Network veracity and thus mental model accuracy are believed to be improved by multimedia designs tailored for dual channel encoding.

Given the recognized and measured limitations of working memory space, Mayer's theory focused upon design decisions aimed at reducing cognitive load demand on the learner in working memory when coordinating information from multiple representations in multimedia and/or computer-based learning. Mayer's theory includes three multimedia principles, modality, contiguity and coherence, important to information presentation and delivery to maximize learner knowledge acquisition (Mayer et al. 2004). The modality principle is based on Baddeley, Chandler and Swell's work describing dual processing channels and builds from the premise that there is an inherent, optimal mode for the delivery of different types (or modes) of information. For example, a visual display of text is less effective and requires more cognitive processing than text presented in an auditory or narrative mode given the method in which this mode of information is most frequently encoded. This principle implies that multiple representations should be constructed so that audio and visual channel inputs occur simultaneously and that the information in each channel is best represented by their particular delivery mode.

Mayer's contiguity principle addresses two separate issues of proximity, both spatial and temporal. Prior research demonstrated that cognitive processing is maximized when related representations of information are displayed in close proximity on the viewing device (i.e. monitor or image) thereby reducing extraneous cognitive load. Learning, problem-solving and the ability to transfer gained knowledge to novel but related situations were shown to be positively impacted by spatially contiguous representations (Moreno & Mayer 1999). Temporal contiguity draws upon dual channel processing in that working memory operations and resulting understanding were shown to be optimized when text and diagrams addressing the same concept were synchronized

21

during presentation. Synchronization enabled working memory to process information sources simultaneously thereby allowing linkages to be constructed between the information representations and thus foster deeper learning and network veracity in long term memory.

The final principle, coherence, or as Mayer and others have more recently termed the phenomena, redundancy, may also be described as information parsimony. When designing learning environments incorporating multiple representations, extraneous (or redundant) information should be omitted resulting in an abstraction of a physical system so that only the most pertinent, causal relationships are highlighted. Informational elaborations should be minimized to include only the specific details required for student understanding. This principle implies that the "bells and whistles" often accompanying multimedia presentations increase extraneous cognitive load, clogging processing channels and working memory, therefore increasing difficulty in student understanding and learning.

These understandings related to cognitive processing and design form the support for the hypothesis that animation delivery will facilitate student understanding. The inclusion of the animations in the treatment groups is proposed to enable these learners to examine an additional representation of the target content and process the information through both the auditory and visual channels. Therefore wind flows and pressure systems will be portrayed dynamically in geographic space affording an additional mode of representation for learners to incorporate into their knowledge acquisition and construction. The control group in this study will not view animations related to the

content area and are expected to perform worse on assessments of understanding and learning gains due to the omission of these representations.

Mayer et al. (2003) also provide support for why the third treatment group in this study (i.e., user-determined animation viewing) is expected to perform better on posttest knowledge assessments and overall learning gains than the remaining three groups. This study found that self-explanation prompting prior to the intervention increased understanding. The examination of text and diagrams while viewing the animations may prompt learner questions thereby enabling the learners to metacognitively monitor disequilibrium between current understandings from text and diagrams and what is being seen and explained in the animation. Thus the group participants can identify the problem area or areas in their understanding and re-visit the specific section in the text, diagram, animation or any of the above representations to seek clarification. While Mayer's work illuminated cognitive and multimedia design elements to facilitate and improve learning, most of the research occurred in a laboratory setting using pre- and posttests with individual learning interacting with a computer is a short time period unlike authentic formal instructional settings involving classmates and a classroom context.

These limitations were addressed when Mayer and colleagues (Atkinson et al. 2005) extended multimedia research in mathematics learning to an ecologically valid setting at both the undergraduate and high school level. Moreover, pretest - posttest assessments were supplemented by including process data indicating users' perceptions on understanding and material difficulty. The study contrasted learning between animations with and without vocalizations and between human and computer generated vocalizations. Human vocalizations were found to have the greatest impact on

understanding. The animations used in this study include human vocalizations describing processes temporally aligned to their symbolic representation. Each treatment group with the ability to examine the animations, with the exception of the control group, will view and hear these animations therefore leading to the expectation of increased understanding.

Section 7. Animation Model of Learning to Inform Design

Wide-ranging empirical comparisons between learning with animation versus learning by alternative representations were conducted by Hegarty and Narayanan from the perspective that animation should improve understanding (1998). As prior work found no benefit to animation inclusion, a design model was formulated to define the cognitive requirements necessary to construct a dynamic model of a mechanical system via mental animation. Six stages were offered. Stage one stated that individuals decompose the machine's diagram due to limitations in working memory storage and prior knowledge. Therefore the diagram presented must be comprehensible and not include ambiguous symbology or text inconsistent with the verbal description provided for the system. Multiple diagrammatic representations are recommended including explosion views, and deictic interfaces linking a component's explanation in the text to the appropriate location on the diagram. Once the system's individual components are understood, the learner can advance to stage two and begin to construct a static, internal mental model of the machine.

In this stage, the individuals seek to build a mental representation of the device by linking the displayed schematic diagram to the real world device and pre-existing

knowledge of the system. Successful linkages allowed the learner to infer component composition and operating functions (and therefore likely behavior when the model is animated in stage five) from the mental representation. Hegarty and Narayanan believed that actual images of the components of the machine are better than diagrams while building this internal representation and should also be linked to explanatory text in order to facilitate the most accurate mental representation. The supposition of diagrammatic or animation realism was contradicted in subsequent research conducted by Lowe (2005).

Within stage two of the model, the learner also must encode the spatial relationships between components of the machine, which are often confused in two-dimensional diagrams. Three-dimensional diagrams with multiple perspectives (or viewing angles) and cross sections clarify component configurations and further add detail and realism to the learner's emerging mental representation. Stage three requires that the learner integrate their evolving mental model with the textual and visual external representations provided in the hypermedia environment. Cyclic iterations and interactions between stages two and three enables the learner to transfer the encoded information from working memory to long term memory with adequate representational detail and linkages to pre-existing knowledge. The hypermedia environment should facilitate mental model construction by co-referencing text and visuals in the same view space, as understood as Mayer's (2001) spatial contiguity principle, related to the system's components. In stage four, the learner must now incorporate the pathways of motion as they propagate through the components of the machine. The causal chain must be identified and linked to the mental model constructed in stages one through three using spatial configurations and prior knowledge. When working with cyclic or advanced

systems, the causal chain may branch and merge thereby increasing complexity and necessitating a sequential explanation of the kinematic processes. The model developers recommend an integrated audio, visual and textual narrative describing the kinematic sequence to encode the proper movements and interactions between components as informed by Baddeley's dual channel research (1992) and Mayer's temporal contiguity research (2001).

When the learner reaches stage five, the mental model is completed with the addition of simultaneous machine movement through mental animation. The movement sequence may be explained by the construction of production rules related to the interactions of each component or by an imagery-based mental simulation of the system, both difficult to access using traditional experimental product data in the form of post-tests. This frequent shortcoming in cognition research related to multimedia learning was addressed subsequently by researchers with the methodological inclusion and collection of procedural data, often in the form think alouds (Azevedo et al. 2004; Chi & Van Lehn 1991; Ploetzner et al. 2005) However, Hegarty's prior mental animation research has shown that novice learners are unable to construct production rules and often reach working memory capacity by having to spatially represent each component's kinematics resulting in longer animation sequences. Moreover, Kozhevnikov and Hegarty (2001) determined in a subsequent study that novices and even experts in the mechanical domain apply inaccurate intuitive knowledge to analyses of movement when performing motion verification tasks. Thus after implementing the model and examining learner performance, a sixth stage in the model was proposed. In this stage, the hypermedia

environment would contain domain specific content tailored to the physical system under examination to address misconceptions found in both novices and experts.

The identification and recognition of misconceptions in learners guided Hegarty, as well as others future research and design recommendations in multimedia toward teaching and learning through conceptual change, a theory described later. However, these early studies continued to believe that the inherent advantages of hypermedia and animation would enable learners to view the kinematic sequence and correct inaccurate representations in their mental model simply with the addition of a content component to be referenced as needed. The inherent advantages noted included animation speed, view angle and whether the motion was displayed sequentially or concurrently through the model. Hegarty and Narayanan believed that repeated viewing worked to correct stage four misrepresentations as did a narrative sequence concurrent with the animation. Designed controls, with the ability to pause, slow, repeat and reverse the animation also allowed the novice learner to stay within the confines of their working memory capacity and modify their mental model. Additionally, the learner may choose to limit the number of representations experienced in the hypermedia environment based upon prior knowledge and expertise or move back down the stages to fill in knowledge gaps, all strategies within the hypermedia environment advantageous to mental model and animation development.

Subsequent empirical studies compared the hypermedia manual whose design was informed by the model against "traditional" paper-based instructional manuals (Hegarty et al. 2003; Hegarty et al. 2001; Hegarty et. al. 1999). The hypermedia model was compared against two types of printed manuals: a complete copy of the information

contained within the hypermedia manual and a second printed manual containing an initial schematic diagrams of the system, a description of the causal chain of motion within the system, and an explanation of the relevant physics principles. The experiment assessed two measures: time required to study the materials and subject comprehension through mental animation multiple choice and open-ended questions (Hegarty et. al. 1999). The major difference between the printed manuals and the hypermedia model were the hyperlinks and animation sequences available in the computer environment. The random assignment of students into the three experimental groups was tested via a background questionnaire containing questions concerning standardized test scores (e.g., SAT verbal and quantitative), prior knowledge in mechanics or physics and practical knowledge related to the topic evaluated through home repair questions. Additionally, a novel measure evaluating subject interest was also included to address motivation and affect. This measure and its potential to impact experiment results were also selected for inclusion in this study.

Results indicated that there was no significant difference in material comprehension between the three instructional environments although significantly more time was spent navigating and processing the hypermedia manual. Hegarty and colleagues stated that the increased study time was a by-product of the subject having to listen to the auditory descriptions of system behavior and having to view a fixed-length animation. However, the hypermedia's failure to increase understanding was more difficult to explain; the researchers offered that perhaps the system, a flushing toilet, was familiar enough to the sample population (i.e., undergraduate college students) that it was unnecessary to provide an extensive explanation of the system. Therefore, the second

experiment augmented the flushing cistern with two additional mechanical systems, a bicycle air pump and a car's brake system, and replicated the first experiment although without including hypermedia manuals. This instructional item was replaced with a simple labeled diagram of the desired system. Interestingly, comprehension results were the same with one notable exception, the air pump. Text describing the causal chain improved comprehension of the air pump as compared to simply examining diagrams with labels. Hegarty and colleagues hypothesized that this difference may be because air pressure and volumes changes are not visible attributes. This observation, relevant to the subject matter used in this research project, led Hegarty and other researchers, Lowe in particular, to continuing research in the domain of atmospheric science given the prevalence of invisible attributes and processes.

The previous study failed to provide an advantage for employing Hegarty's cognitively informed hypermedia model when compared to printed materials containing the same content. Therefore Hegarty and colleagues selected another approach and attempted to determine whether the model improved comprehension as compared to commercially produced CD-ROMs examining mechanical reasoning (i.e., David Macaulay's The Way Things Work) (Hegarty et al. 2001). Subjects using the theoretically designed hypermedia manual performed better on measures of comprehension as compared to viewers of the CD-ROM. However, additional experiments reinforced results from prior research showing no multimedia advantage. Comprehension differences were not observed in comparisons of text and hypermedia treatments of the same topic using the same design format. In other words, the format was unimportant. This finding was further supported and elaborated upon in Narayanan

and Hegarty's 2002 study (Narayanan & Hegarty 2002). Printed and hypermedia materials constructed using their theoretical model increased learner comprehension when compared to books and CD-ROMS leading the researcher's to conclude that content and structure are more important than format. These results led the researchers to hypothesize that perhaps the systems examined in their current research were simple enough for subject's to mentally animate solely from a single, well constructed diagram. Therefore abstract and more complex physical systems containing invisible components (e.g., air motion) were introduced in subsequent studies to evaluate the benefit of hypermedia presentation based on their theoretical model.

Hegarty et al. (2003) examined the impact of external animations on the mental animation process and found that the internal mental animation process is more important than the external representations, if equivalent information is conveyed in each. Hegarty concluded that perhaps static representations offer advantages in that diagrammatic simplifications and abstractions present only the most relevant features of the material under consideration, an idea Lowe (2005) simultaneously advocated and explored in his salience research with meteorological maps and animations. Hegarty stated that critical processing time and space in working memory are not consumed by encoding confusing and irrelevant information with complex animations (Hegarty 2004). Two additional observations were reported concerning animation use for learning at the fundamental level. First Hegarty offered that perhaps the cognitive processes involved in encoding and processing differ based on the mode of presentation. Viewing an animation is inherently more passive (the "couch potato" phenomena) than viewing a static diagram and because rather than inferring motion, the animation provides visualizations of the system's

motion. Therefore, the attention offered the content through executive control is unequal

and can result in less comprehension given the dynamic presentation. Secondly, Hegarty

stated that perhaps the nature of the content under study and the specific learning goals

measured by the experimental assessments resulted in comparative disadvantages when

the material was presented dynamically (Hegarty 2004). For example, Hegarty asked

whether animations afforded the same instructional advantages when used to present

chemical and/or meteorological phenomena as compared to mechanical phenomena.

Unlike the simple representations used by Hegarty, most models of real world

physical systems are complex and include translations and rotations in four-dimensions.

Moreover, these environmental systems contain components that are not visible, for

example horizontal and vertical atmospheric gas movements (or wind) and occur at scales

beyond the human scope (e.g., hundreds and/or thousands of miles). The recognition that

the animation medium was not inherently better than static diagrams for some domain

content guided subsequent researchers to examine different domains and to focus upon

how learners constructed meaning and understanding from animations through design

decisions. While the animations used in this research project are not testing design

differences, textbook animation creators have incorporated design recommendations from

the literature especially related to user control and functionality and the inclusion of

narration.

Section 8. Animation in Learning Interactivity and Design

Research has shown that to improve learning in multimedia environments,

students must be able to extract the relevant information from the representation using

prior knowledge of the subject area as well as contextual cues provided by the interface's design (Lowe 2003; Mayer 2001). Interactivity and active participation with animation representations is recommended by designers to improve learning from the constructivist paradigm (Bodamer 2004; Goldman 2003; Lowe 2003, 2004; Mayer 2001). Relevant information is encoded by the learner, organized, and incorporated into the student's existing schema or mental model thereby inducing conceptual change and resulting in increased understanding. Moreover, empirical research has shown that the environment must be designed to highlight fundamental concepts prior to the presentation of complex scientific systems (Bodamer 2004). Initial concepts must be presented at the user's pace (implying user control) due to differences in prior knowledge and working memory capacity while also providing scaffolding to avoid the propagation of misconceptions. The presentation of dynamic visualizations or animations, therefore, is only recommended after sufficient background knowledge exists and only when the mode fits the content (Goldman 2003). Lowe (2004) produced a similar recommendation concerning prior knowledge. Both Goldman's and Lowe's recommendations directly inform the research questions addressed in this study.

Goldman (2003) suggests animation use should be limited to content that matches the presentation mode therefore the target content must be inherently dynamic in nature. Given the constant dynamism of the earth's atmosphere due to thermal differences on the surface, the content area selected for this study unequivocally aligns to this recommendation. Secondly, Goldman (2003) and Lowe (2004) suggest that learners must have some threshold level of prior knowledge in order to extract and encode the most relevant concepts and interrelations portrayed in the animation presentation. This work

suggests that participants of the treatment condition defined in this study where the animations are viewed prior to reading the text materials should perform at a lower level on the posttest assessment tool than the two treatment conditions when the animations are viewed after or concurrent to the text reading. This supposition assumes that reading the text prior to viewing the animations provides sufficient prior knowledge to selectively encode the relevant phenomena represented in the animations.

The importance of prior knowledge acquisition is noted to reduce the maintenance and/or propagation of misconceptions. Misconceptions can be understood as incomplete or inaccurate mental representations of the content area confounded by limited prior knowledge. These mental representations or causal mental models are theorized to exist in long-term memory and enable the learner to reason about physical systems when given an external representation similar to a prior experience (Markman & Gentner 2001; Hegarty 1992). Mental models provide a qualitative method for the learner to reason and understand physical system functions however, if prior knowledge is limited, these models are often surficial and fragmented (Markman & Gentner 2001).

Section 9. Learning from Animation vs. Static Images

Empirical research by Lowe (2003) illuminated one possible explanation for mental model gaps by examining how students sought to understand information presented in an animation sequence. Learners given control of the animation were shown to examine the sequence in segments and to encode features that varied in space and time (i.e. Mayer's spatial and temporal contiguity or the dynamic portions of the animation), disregarding relevant static, contextual information (Lowe 2003). Complex systems, as

represented by the content of these animations, entail multiple cause and effect loops at various scales significantly hindering novice learners' knowledge acquisition without sufficient guidance or scaffolding (Jacobson & Wilensky 2006; Lowe 2004). To effectively comprehend the content of the animations, learners need to metacognitively monitor both the acquisition of new knowledge from the dynamic learning environment and accommodate this information in their preexisting long-term memory schema (or mental model) for conceptual change to occur (Hmelo-Silver & Azevedo 2006). These monitoring skills are often underdeveloped in domain novices thereby limiting the ability for conceptual change to occur when viewing an animation (Azevedo et al. 2004).

These findings support the contention that participants in the animation prior to reading group will exhibit a lower level of mental model development given their lack of knowledge about the specific target content thereby inhibiting their ability to metacognitively monitor understanding during the learning episode. Learners who have read the textual description prior to animation viewing should be able to monitor their understanding to a greater extent given increased system familiarity and therefore produce greater increases in mental model levels.

Lowe's research is informative to this study given his frequent focus upon the meteorological content area and his work in detailing how learners interact with animations during learning episodes. Lowe (2008) approaches the use of animation for conceptual understanding from the model developed by Narayanan and Hegarty (1998) whereby three potential sources for learning bottlenecks are discerned given the mismatch between animation information delivery and learners' processing ability. These bottlenecks include the issue of concurrency, speed and complexity. Concurrency

problems in animation result from design decisions where all system movements are displayed simultaneously rather than sequentially counter to Hegarty's finding that individuals deconstruct animations and construct mental models in a serial fashion (Hegarty 1992; Narayanan & Hegarty 2002). Moreover, the speed or pace at which the animation is delivered is often faster than the learner's working memory capacity resulting in incomplete understandings and a fragmented mental model. This issue can be exacerbated by the complexity of the animation presented in the given time period however can also be mitigated if user pace controls are embedded in the software. Designers tend to construct animations based on technological advances in computational power and graphics delivery without regard to the limitations of novice learners' prior knowledge base and working memory capacities including the spatial abilities of mental rotation and transformations common in four dimensional representations. The transient nature of the representations coupled with these limitations may inhibit the development of "runnable" mental models. While the animations used in this study do not allow students to alter the timing of the depicted events (i.e., frames per second), the animations' design enables the viewer to re-examine the target content as frequently as the viewer desires by automatically repeating the sequences until the user selects to the stop button. Moreover, the animations sequentially add system components before depicting the entire interrelated system reducing the impact of concurrency issues described by Lowe. Moreover, when the animations are restarted by the learner, the depicted circulation system is removed entirely from the frame of reference and reconstructed in its entirety from the blank state.

Lowe's work on learning from animated meteorology maps found that even if pace control is modifiable by the learner, understanding did not significantly increase (Lowe 2003). Domain novices were found to spend more time on the initial state (or frame) presented by the animation and the final state (or frame) when seeking to respond to questions answered by the animations neglecting the "meat in the middle" (Lowe 2008). Additionally, learners were found to encode perpetually salient seductive details (i.e. individual components) embedded within the representation at the expense of developing a higher order understanding of total system functioning or the big picture perspective. Thus their resulting mental models did not enable learners to actively predict future atmospheric states or infer conditions or interrelationships between displayed components. Given the collection of data on dwell times of animation frames and the controls users chose to employ, Lowe concluded, dovetailing Hegarty's early work, that animation learners made the dynamic representation static thereby eliminating one of the perceived advantages of the medium. Lowe states that without training and/or guidance, learners are unable to extract complete system and subsystem functioning given the focus on parts rather than the whole and the choice of trying to infer system operations from fixed states calling into question the utility of presenting novices with complex animations. These issues are not as relevant to the atmospheric processes depicted in the animations used in this study given the continuous nature of cyclonic and anti-cyclonic movements and Hadley Cell circulation in which there is no beginning or end.

Lowe continued to examine the linkages between static diagrams and animations in concurrent research at the theoretical and empirical level. With Schnotz, Lowe argued that treating the two representations forms as distinctly different is a mistake given that

our cognitive system operates on both in the same manner given the evolution of the system through interactions with static and dynamic environmental inputs (Schnotz & Lowe 2008). The authors report that representational realism and aesthetics drive design, including animations, even though visual communication is shown to be facilitated when specific features of depicted content are emphasized employing elimination or simplification techniques. Schnotz and Lowe (2008) argue that animation design should focus upon the inherent advantages of presenting system functioning in a non-realistic manner but in line with how humans encode information. Recommended design components include building hierarchy into the animation so that complex parts or functions are constructed at different levels of granularity displaying structures and interactions explicitly through exploded views. Temporal realism would thus be eliminated to the benefit of behavioral realism in that causal relationships between system components would be exhibited and explained. By directing learner attention through the causal sequence, relevant areas or events should be highlighted and made salient enhancing cognitive processing. Research has shown that animations provide different advantages for learners based on their level of prior knowledge. Schnotz and Rasch (2008) find that in high prior knowledge learners, the animations serve an enabling function allowing these learners to perform more complex tasks because the animation allows working memory to offload cognitive load leaving more processing power for generative load. However, in low prior knowledge learners, the animations are shown to serve a facilitating role as the animation enabled the learner to construct the mental simulation of what was displayed. While this learner would never reach the same level of understanding as the high prior knowledge learner, the animation was nonetheless

beneficial. This finding has direct implications to results presented in the current study. Because the animated system and content area selected for the study is fairly novel to the participants, treatment conditions viewing these animations should be expected to develop a more robust mental model in comparison with the control group (i.e., not viewing the animations) given the facilitating function of the representation.

The authors next addressed spatial and temporal dynamics in animations and how these advantages should be employed for maximum impact. Because humans' cognitive systems are designed to recognize motion and pattern, these abilities enable prediction to future states in time. Thus the authors recommend suppressing spatial and temporal invariant aspects of animations in favor of contrasts so that relevant processes are adequately highlighted against static backdrops. This idea led to the additional recommendation of temporal categorization within the animation. Designers were advised to eliminate repetitive, non-salient frames in sequences and focus upon time periods or key frames when state changes occurred in the target system. Thus a parsimonious chunking of the target content was theorized to deconstruct macro-events into the critical micro-events focusing on the transformational time periods. This temporal structuring offloads redundant working memory processing in the learner allowing for attention to the key events for system understanding. Static diagrams have been utilized in much the same way when a series of frames is presented to represent spatial and temporal change. Each frame or time period in the animation is selected for highlighting based on the quantity or importance of change taking place at that moment. The goal is to enable learners to perceive these changes at key times given the selectivity of visual perception in our cognitive systems. In other words, most learners cannot attend

to all elements in an animations sequence at the same level. In static diagrams, given their non-transient nature, learners can interrogate the depiction for understanding with no time constraints and attempt to encode all displayed relationships. In contrast, animations change in time, thus visual perception must be guided to areas and/or events to maximize encoding. Research has shown that novice learners enter into animation perception in a bottom up direction where the salience of the feature or process guides attention whereas domain experts examine the animation from a top down perspective. Prior knowledge in the content area allows the expert to ignore irrelevant symbols or locations in the animation while focusing on the processes that will move the learner toward their learning goal. Experts enter with a specific set of problem-solving strategies generated from experience with the content area and seek to reach the goal of the exercise often an answer to a question. The novice attempts to encode conspicuous information only without the benefit of guided selectivity exhibited by the expert.

The solution offered by the researchers is to employ controls in design that align the animation presentation to learners' cognitive processing ability. While users may be able to control delivery speed, the animation should stop at process transitions on key frames and signal critical events by highlighting with arrows, text and/or flashes to provide a top down hierarchy that novice learner's lack. Thus a balance between free exploration and guided interactions must be programmed given the goal or learning outcome embedded in the animation rather than design based on current technology. User interaction is identified as critical so that active learning is initiated rather than the animation being viewed passively as an entertainment device.

Goldman (2008), while recommending the same design decisions advocated by Schnotz and Lowe, asks larger questions of animation use for learning by acknowledging individual differences in learners' prior knowledge level, attentional capacity and spatial abilities. Thus animations serve each group in different capacities and should be designed to facilitate learning for each group's strengths and weaknesses. She also states that animations may not be appropriate for all content areas and learner groups. An example is provided that learners often create visuals to aid their understanding of topics but do not create animations (However this may change with technological advances in personal computing. Currently this is impossible with paper and pencil and would be a time consuming task if programming on an available laptop). Most often static diagrams are created as learners tend to generate what they've seen in textbooks or have been exposed to during instruction. Moreover, Goldman recognizes that the affordances offered by animations require a level of prior knowledge that some students may not have. If novices do not understand the component parts of the animation, the dynamic interrelationships displayed will be beyond their zone of proximal development thereby inhibiting any potential learning. Goldman observes that most animations are made from an expert's vantage point and may not provide what is meaningful and important to the target audience. These recognitions are critical to textbook publishers' animation developers and have been noted to guide changes in included animations in the last decade. While the animations used in this study do not include exploded views or salience signaling, representational simplification recommendations and sequencing have been observed.

Ploetzner et al. (2008) present a similar perspective in noting that the quantity of information contained in animations often overburdens learners. These researchers promote the explicit instruction of learner strategies to facilitate understanding from dynamic visualizations at all levels of schooling if animation learning is to be successful. An experiment was conducted where students transformed information between text and graphics and then attempted to solve a computer puzzle. Results indicated that successful learners applied more strategies, for example making drawing and taking notes, and were less wary about proposing solutions even though they knew their answers were not entirely correct. Less successful learners did not try to generate a solution state and re-visited the learning materials any time difficulty was experienced. Ploetzner et al. concluded that the less successful learners were encoding the dynamic information at a superficial level and had difficulty coordinating the text with the representations.

Section 10. Product and Process Data in Animation Research: Mental Models

Assessments to evaluate understanding in animation research span the continuum from product data including multiple choice responses eliciting declarative knowledge to open-ended and diagram questions allowing for less discrete and more robust knowledge representations. In addition, researchers interested in the cognitive processes employed by learners in multimedia and animation environments have also collected process data in the form of eye tracking measures and learner vocalizations to provide a more comprehensive inventory of learner artifacts. This study employs both product and process data collection to shed light on separate features of the experiment with the use of mental model levels to examine conceptual change occurring in each experimental group.

Mental models provide a powerful framework for education researchers to explore understanding and learning. The concept of a mental model was introduced to explain human thinking and reasoning about the natural world by Kenneth Craik in 1943 (Kaplan & Black 2003). Since its introduction, educators, psychologists, cognitive scientists, and physical scientists interested in the function of the human mind have sought to illuminate the development and structure of human thinking by delving into the nature of mental representations and their impact on understanding. Considerable research in these disciplines, employing a variety of methodologies, has extended the initial conception of a mental model to a fine- grained model that includes components, component linkages, model applications, and evolution (or conceptual change). Mental models are operationally defined to mean internally constructed representations of natural phenomena (Gentner & Stevens 1991) or more specifically the process by which humans' model complex systems in the physical world to generate inferences and predictions (Clement & Steinberg 2002). Most research incorporating mental models related to learning and understanding is from the domain of science knowledge (Albert1991; Clement & Steinberg 2002; Gentner & Gentner 1991; Lehrer & Schauble 1998; Mayer, Dyck, & Cook 1984).

Mental models represent human understandings of phenomena of the natural world and are based on observations and interactions with it (Gentner & Stevens 1991). These models are theorized to form through everyday experiences and are employed in an attempt to understand dynamic processes by simplifying the complexities of the world into a discreet number of causal connections (Albert 1991; Gentner & Stevens 1991). This simplification process led Johnson-Laird (1983) to posit that "there is no complete

mental model for any empirical phenomena." Given the models' dynamic or "runnable" nature, mental representations enable individuals to generate inference and prediction about systems using their personal cognitive structures (Vosniadou & Brewer 1992). Moreover, another dimension of the dynamic nature of mental models is the ability for these constructs to be continuously modified by interactions with the natural and human environment and through instruction (Norman 1991). Frequently, these models attempt to represent phenomena and processes that are not directly visible or quantifiable resulting in qualitative representations (Taylor, Barker, & Jones 2003). However, as individuals seek to become experts in the domain, Lehrer and Schauble (2003; 1998) state that model robustness and voracity is improved when individuals seek to "mathematize" their understandings of a system by transforming their mental representations into a more symbolic and mathematical form.

Cognitive scientists further specify that qualitative mental models represent declarative, procedural, and inferential knowledge in an attempt to represent complex (e.g., science) topics (Greene & Azevedo 2007). The inferential nature of the mental model is thought to result from the generative capabilities of human cognition as individuals seek to connect disparate fragments of knowledge about the phenomena through linkages (Vosniadou & Brewer 1992). Theoretical representations of the componentry of mental models often place declarative knowledge in nodes and represent procedural knowledge as linkages between these nodes. Thus a mental model or internal representation is built incrementally as individual components of the system under study are represented (or learned) and the linkages between nodes or knowledge objects are established to convey causal relationships (Mayer & Chandler 2001). However,

knowledge nodes and linkages incorporated into the mental model construct are found to vary considerably between individuals seeking to understand the same phenomena and are far from unbiased (Libarkin, Beilfuss, & Kurdziel 2003).

Researchers employing the sociocultural constructivist lens contend that these models develop through interactions and discourse with others as well as through direct instruction in a social context (i.e., classroom) therefore conveying epistemological and ontological beliefs from the instructor and societal norms that shape the entire structure of the individual's mental model (Col & Treagust 2003; Human-Vogel 2006). Therefore, pre-instructional interaction in the realm of physical system learning often leads to the inclusion of preconceptions, misconceptions, and alternative conceptions (i.e. naïve intuitions) which are frequently contradictory to the manner in which the system is understood in the scientific community (Vosniadou & Brewer 1992; Williamson & Abraham 1995). Interestingly, given these considerations, mental model researchers often state that the purpose of a person's mental model is to provide a construct to enable sense making and reasoning (in a scientific manner) even though significant barriers to correct inference and prediction exist (Kaplan & Black 2003; Merrill 2001).

Norman (1991) and Johnson-Laird (1983) state that mental models are by their very nature incomplete and at their worst incorrect in representing the phenomena under study. Moreover, model details are forgotten with limited use and models tend to blend together when employed in novel learning domains and situations (Ke, Monk, & Duschl 2005). Even following instruction, individuals continue to maintain "superstitious," naïve, and unscientific beliefs in an effort to minimize mental effort and to maintain model parsimony (Clement & Steinberg 2003: Norman 1991). Conceptual change (or

mental model evolution) in a complex science domain (e.g., quantum physics), even given years of instruction, was discovered to take place incrementally as certain linkages weakened and others strengthened with components of prior mental models always retained (Ke et al. 2005). Ke's finding supports evidence that mental model representations that do not align to canonical science (i.e., alternative conceptions) can detrimentally impact and hinder conceptual understanding during instruction.

Libarkin et al. (2003) present a categorical system for mental model understanding in an attempt to synthesize cognitive scientists' knowledge on human understanding related to science learning. Four categories of cognitive models were developed to represent the continuum from novice learner to expert in the domain. Their first category, naïve mental models, is based on individual's conscious and unconscious observations in the natural world and the commonsense intuitions resulting from repeated interactions and experience with physical phenomena. These models are described as general, unconnected, and fragmental bits (or nodes) of knowledge which are equivalent to diSessa's (1993) phenomenological primitives or p-prims. The authors state that these models are spontaneously created in an attempt to understand novel situations or new information. The second category of mental model is described as an unstable mental model, due to the fluid or highly modifiable nature of the cognitive construct. Given the incomplete nature (or gaps) of the individual's understanding about a phenomena, new information reorganizes the existing explanatory model as linkages between the p-prims are developed or severed. After repeated interactions and direct experience with the phenomena, the mental model becomes more organized and stable given fewer and fewer conflicting observations. This mental model category is described as a conceptual

framework and is frequently employed by the individual to comprehend physical interactions. Each of the three stages described above are considered novice mental models due to their internal, personal nature and the lack of formal instruction or inspection from an expert community in the domain. The fourth mental model, described as a conceptual model, is the model associated with expertise in the domain. Libarkin et al. (2003) describe the model as the representation developed and used by the domain experts and is therefore external or communicated in a written and oral form for inspection. This model is highly stable and precisely defined, either mathematically, analogically, or via physical models. Due to the model's precision and robustness, changes occur only following significant effort and time as contradictory evidence emerges and alternative explanations are complied, debated, and eventually accepted. Given the external and communicative form of the final model, this understanding is stated to be "accessible to any individual" alluding to the unstated implication that instruction can bridge the gap (i.e. accomplish conceptual change) between an individual's conceptual framework and the scientists' conceptual (or scientific) model.

In many cases, these instructional experiments and interventions use the conceptual or scientific model as the goal for learners' understanding through instruction although developmental levels (i.e., student ages) dictate the precision and level of detail expected within the target model. Thus the scalable nature of the conceptual model often results in the teaching of functional analogies to represent the system's behaviors as closely as possible. Thus, assessment levels must align to instructional models.

Section 11. Instructional Design for Conceptual Change

Conceptual change assessment based on external (i.e., graphically depicted, vocalized or written) mental models has been conducted on student understanding of scientific concepts and systems across the content areas of physics, chemistry, biology, and the geosciences using students at nearly every level of education. A recent review of this research work has illustrated that eclectic methodologies have been employed resulting in a broad spectrum of qualitative and quantitative tools each attempting to characterize knowledge gains given an instructional intervention (Gentner 1991).

Lehrer and Schaube (1998) compared the modeling capabilities and differences of 2nd graders with 5th graders' in an effort to understand causal reasoning about structure and form as related to gear interactions in physical models. The authors described two types of explanatory model structures, mathematical and mechanistic, necessary to reason about gears in a scientific or model-based form as compared to the naïve physics understanding based solely on observation. Open-ended questions were used to determine the level of understanding related to model variables including the transfer of motion, the direction of motion, the speed of gear turning, and the understanding of mechanical advantage. Data were collected by audiotape and videotape (to assess the hand motions employed by the children) during the course of two interviews as the children interacted with the gears and gear based models. For each model variable, categories were developed along the continuum of no understanding to sophisticated understanding based on the children's explanations. Results indicated that both grades sought to explain the experiments using causal theory, however only 5th graders were able construct the causal chain in complex configurations. Neither group was found to employ mental models

using model-based mathematical and mechanistic reasoning because they had never been asked to express these relationships symbolically or defend their understandings through discourse within the learning community.

Following the analysis of empirical research in children's conceptual understanding of science processes, Lehrer and Schaube (2003) refined and expanded their theory of model-based reasoning to inform an instructional strategy to scaffold the gap between naïve physics understanding (i.e., unstable mental models) and conceptual models or scientific understanding. The foundational key for the proposed instruction was to explicitly introduce models and modeling because these constructs are one of the most important and authentic enterprises undertaken by practitioners of science. Moreover, Lehrer and Schaube (2003) state that the "modeling game" provides an effective method to convey the symbiotic relationship between mathematics and science and, when implemented throughout the science curriculum, enables deeper model understanding and realistic forays into science's understanding of the physical world. Model-based reasoning was proposed to enable multiple, beneficial forms of representational mapping in young children (i.e. learners). For example, symbolism and analogy use as representational components of models were exemplified as structural improvements to systemic understanding.

As models became more deeply understood and sophisticated, learners began to evaluate representational choices systematically by comparing alternative and rival models and seeking to discover missing or erroneous components. In this process, the models become less concrete (one to one direct relationships between the model components and processes) and more abstract especially when symbolic representation

takes the form of mathematical explanations. "Mathematization" was stated to increase model transport and extension across physical processes noted in the real world (i.e., transfer) and aided in solidifying (or quantifying) individual's assessments of their personal mental model and alternative or competing expressions. Lehrer and Schaube's (2003) methodology informed the development of the ordinal mental model scale developed for this study. Levels were assigned on a continuum from no knowledge (or naïve system understanding) to expert system which was defined as replicating the model presented in the text, graphics and animations.

Coll and Treagust (2003) examined the impact of teaching models on students' conceptualizations of ionic bonding in secondary, college, and graduate students and the changing nature of the students' mental models with additional instruction in the field. Teaching models, used to shape student's mental models, had been defined as expressed models that have the characteristics of completeness, coherence, concreteness and correctness. However, the identification and explanation of the models and their components and functions were found to vary across learning environments. Coll and Treagust explained that these models were expressions of consensual understanding contextualized to the institution in which they were used; however, being models themselves, these representations were imperfect and simplifications often containing erroneous facts. Because of these issues, the researchers recommended that model limitations be explicitly presented to learners during instruction in order to deconstruct the model as truth myth. Additionally, students' alternative conceptions resulting from prior experiences, their sociocultural context and exposure to conflicting teaching models (i.e., text model differs from instructors' expressed model) impacted the mental

constructs evaluated during research (Coll & Treagust 2003). Their findings suggested that instructors understand and address the difficulty that students have in releasing pre-existing beliefs and that presented models should be taught at the appropriate level.

Hogan, Natasi, and Pressley (2000) presented an instructional sequence that incorporated the expression of alternative conceptions prior to the presentation of the scientific communities' accepted conceptual model in their analysis of how discourse can induce mental model change. In their model, teachers served the critical role of requiring that student clarified and crystallized their thinking through Socratic questioning. Their four phases began with the expression of students' mental models (i.e., current conceptions), the construction of a mental model given experiments to produce coherent explanations and predictions, the verbalization of their evolving mental models for discussion and refinements after their presentation to the class, and the use of the newly constructed model to explain new observations. While this research did not adopt a direct instruction technique and therefore employ Hogan et al.'s model, having students diagrammatically represent the target model prior to the learning episode served to illuminate tropical weather and climate misconceptions identified and discussed in the analysis section of this study.

Taylor, Barker, and Jones (2003) also examined mental model alignment between the learners' representation and the scientific community's conceptual model by providing an instructional framework based on conceptual change in the realm of mental models. The authors' stated that mental models are a "core process in astronomy itself (p. 1206)" and provided examples of how the discipline has been advanced through model evolution and change from Ptolemy to Kepler. The authors examined mental model

development from a sociocultural framework for both scientific practitioners and learners in the discipline. The sense making component of mental models was shown to direct their dynamic evolution and in the context of science, the external expression of these mental representations was said to initiate interrogation, revision, and development as new observations and understandings were incorporated. However, as noted by Coll and Treagust (2003), learners rarely understood what models really represent due to the lack of explicit instruction about their nature. Moreover, these authors stated that learners see model modifications as an error correction process rather than a constant advance of understanding (i.e., learning) given continued inquiry. Taylor et al. (2003) found that student presented mental models may not represent true conceptual change but may be constrained by what they perceive to be the classrooms' social norm or context (Taylor et al. 2003) rather than true modifications and reorganizations. Therefore, assessment of true conceptual change can be difficult.

The design of this research study sought to reduce the impact of social context by having students study and understand the target model independently albeit surrounded by classmates in their usual instructional settings. While this may have impacted their ability to have misconceptions and pre-conceptions modified through discourse and argumentation, the examination of learner's pre-test externally represented mental model with the posttest mental model enabled the researcher to identify specific cases of misconception maintenance in direct opposition to model fidelity.

Section 12. Conceptual Change/Mental Model Assessment Challenges

In many cases, the assessment of learning gains associated with computer animations and interactions was couched within the mental model framework to aid in defining specific criteria for assessment. Williamson and Abraham (1995) directly addressed the impact of animation use in instruction on college students' mental models of particulates in chemistry. Their research discovered that students have the greatest misconceptions about atomic and molecular processes at the microscopic scale and developed detailed visualizations to illuminate these previously unobserved behaviors. Using quantitative assessments based on the agreement between the students' mental model and scientists' conceptual model, the researchers indicated that the animation treatment group's understanding increased relative to the static diagram and text only groups. Their explanation for the increase was based on Pavio's dual coding theory which states that pictures are coded into memory in as both verbal and imaginal codes thereby increasing their recall as compared to verbal representations alone. A similar experiment was conducted in the domain of chemistry using both computer animation as well as an additional visualization tool, video clips (Velazquez-Marcano et al. 2004). Students were shown videotapes and animations of a chemistry experiment examining fluid equilibrium and asked to predict the behavior of the gases or liquids in three experiments. The treatments varied between video first and then animation and animation first then video and predictions were made after each mode of presentation. Students were found to predict the behavior of the fluids more correctly after watching both representations (rather than only one) although the order of the viewing was not significant. Each of these studies implied that mental models were made more sophisticated given the computer-

based dynamic representations and the ability to concretely visualize formally invisible processes. These results support the contention that the three treatment groups should outperform the control group given their ability to view additional representations of the target content.

However, these results have not been uniform in all animation studies across all domains. Mayer (2001) reported that computer-based learning environments do not necessarily increase student comprehension in every case. His research indicated that working memory can become quickly overwhelmed when computer-based instruction is not carefully designed and presents too much content too quickly (Mayer 2001). Thus, he designed an experiment in which users controlled the pace of words, pictures, and animations in the domain of meteorology (Mayer & Chandler 2001). Previous research found that users did not explore as much information when navigating independently as when guided by the environment so this program required users to visit each component in the hypermedia environment in two distinct sequences: one group was required to watch the entire presentation first and then return to examine individual components while the second group was able to watch each component before viewing the entire presentation. The second group scored better on tests for transfer and inference indicating that encoding small quantities of information before trying to comprehend the entire sequence was more effective. These results indicated that mental models are constructed in two stages. In stage one, component functions must be encoded before the linkages between the individual components are incorporated into a more robust model of understanding achievable at stage two. This finding aligns quite closely with Hegarty's piecemeal model for mental animation.

More recently, researchers seeking to more fully develop the methods in which computer-based learning can increase mental model complexity began to incorporate cueing or tutoring routines in the hypermedia environment to aid learners in conceptual changes. Kaplan and Black (2003) described a computer-based hydrology modeling environment in which students sought to determine casual factors related to flood occurrence by varying levels of specific environmental variables (e.g., soil type, soil depth, water temperature) in iterative simulations. These simulations were designed to provide inquiry-based deductive reasoning and were augmented with cues in the form of field reports to aid student recognition of the most relevant factors. Mental model assessments were conducted quantitatively by summarizing the number of model components, linkages, correct inferences, and evidence-based explanations generated during the simulation activity. Results indicated that cues were highly effective and positively impacted the complexity of the mental models created by the learners.

As understandings of conceptual change as evaluated by mental models continued, researchers began to approach conceptual understanding from the perspective of learner employed cognitive processes (Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). To uncover specific, successful strategy use, some researchers relied upon learner generated self reports, with their inherent recall limitations, while others adopted think aloud protocols, enabling the researcher to record and code vocalizations of cognitive actions while actively engaged with the hypermedia or multimedia environment (Winne & Jamieson-Noel, 2002). The most applicable and comprehensive coding strategies in the exploration of self regulatory learning (SRL) behaviors relevant to this research study have been produced by Azevedo and his protégés. The following section

will introduce key concepts of the theory of SRL and describe results from several empirical studies utilizing procedural data collection and analysis. This research study applied Azevedo's SRL coding methodology and analytic framework to illuminate cognitive processes prevalent in the subjects in this study's context.

Section 13. Self Regulated Learning, Procedural Data and Mental Model Representations

Azevedo and colleagues evaluated mental model change in hypermedia environments to determine the specific self regulatory learning (SRL) processes students enact to learn science content through the collection and interpretation of procedural data (Azevedo et al. 2004; Azevedo & Cromley 2004; Green & Azevedo 2007). Procedural data collection seeks to determine the specific cognitive processes enacted when learners coordinate information presented by multiple representations with long term memory to develop or improve their conceptual understanding, or mental model, of a science system topic (e.g., the circulatory system). Self-regulated learning is viewed as active, intentional learning whereby learners establish learning goals and "attempt to monitor, regulate, and control their cognition, motivation, and behavior" in an effort to reach constructed goals (Azevedo & Cromley 2004, p. 523). Azevedo's model is based on Pintrich's (2000) four phases of SRL which include 1) planning and goal setting, 2) monitoring (using metacognitive processes) 3) monitoring learning and motivation and 4) reflection. Intentional learning in empirical studies was parsed into planning, monitoring, strategy use, task difficulty and demand and interest and learners' cognitive actions were coded to discern successful and unsuccessful strategies. An example of the major categories and sub-codes are presented in Table 1 from Azevedo and Cromley (2004).

Table 1. SRL Behaviors from Azevedo and Cromley (2004)

Planning
  Prior knowledge activation
  Planning
  Recycle goal in working memory
  Subgoals

Monitoring
  Feeling of knowing
  Judgment of learning
  Monitoring progress toward goals
  Identify adequacy of information
  Self-questioning
  Content evaluation

Strategy use
  Draw
  Summarization
  Taking notes
  Read notes
  Knowledge elaboration
  Coordinating informational sources
  Find location in environment
  Selecting new informational source
  Goal-directed search
  Free search
  Evaluate content as answer to goal
  Mnemonics
  Inferences
  Rereading
  Hypothesizing
  Read new paragraph
  Memorization

Task difficulty and demands
  Time and effort planning
  Control of context
  Help-seeking behavior
  Expect adequacy of information
  Task difficulty

Interest
  Interest statement

Azevedo's research indicates that learners are often unaware of the specific

cognitive strategies necessary to successfully understand complex scientific systems,

especially in hypermedia or multimedia environments where information is presented

with multiple representational modes. Azevedo and Cromley (2004) tested this

hypothesis by testing the impact of SRL training on learning in comparison with a control

group with no SRL training. Results indicated the SRL trained groups' mental models increased significantly more than the control groups. Moreover, when examining the specific SRL processes undertaken during the learning episode, the experimental group was discovered to use a more diverse and effective set of SRL processes than the control group. For example, this group more frequently referenced prior knowledge in the planning phase, metacognitively monitored their learning repeatedly to ensure content understanding, and deployed effective strategies including note-taking, summarizing and coordinating representations. In contrast, the control group was found to focus on the external environment by recycling goals in planning (i.e., spinning wheels) and frequently querying the information for adequacy versus integrating with prior knowledge.

Concurrently, Azevedo, Cromley and Seibert (2004), sought to produce empirical evidence to support how specific forms of scaffolding are more effective than others again using process or procedural data. In this study, the group compared three conditions, an adaptive scaffold in the form of a human expert to guide student's learning and progress through the content, a fixed scaffold which deconstructed important elements of the target content into a series of sub-goals (i.e., ten questions/statements) to guide students, and a no scaffolding condition in which students deployed their own learning strategies with no external guidance. Procedural data were evaluated through the coding and frequency counts of think alouds for comparison

The study found that the adaptive scaffolding group implemented higher level SRL strategies than the fixed and no scaffolding groups, and displayed the greatest mental model shifts. Surprisingly, Azevedo et al. (2004) report that the adaptive

scaffolding group seemed to off load SRL tasks to the human tutor thereby reducing their overall usage during the learning episode. However, the SRL strategies used by these students tended to be more effective and efficient, for example evaluating understanding by summarization and judgment of learning than less effective strategies such as recycling sub-goals and searching for information haphazardly. These results provided an empirical foundation for the importance of specific SRL processes when coordinating information in a hypermedia environment and were used to argue for student training in SRL processes to improve understanding.

Green & Azevedo (2007) found that certain SRL processes observed in learners while navigating the hypermedia environment significantly improved these individuals' mental models of the observed content when compared to pre-existing conceptions. Students who compared multiple representations of the same information, returned to fundamental concepts while viewing complex interactions, and practiced inference and knowledge elaboration (e.g., analogy use) exhibited significant, qualitative shifts in mental model complexity.

Azevedo's use of mental model development, evaluation and change to measure learning and the collection of procedural data to discern beneficial strategies applied by learners in multi-representational environments have been adopted by subsequent researchers, Butcher (2006) and the author of this study. Azevedo's SRL process data collection methodology and coding classes provide the framework through which this study interprets participants' learning strategies during the learning episode.

Section 14. Literature Review Conclusion

Theories of learning, initially the domain of psychological researchers, entered education as a guiding framework through which educational researchers could interpret the complex and messy environment of formal education. Work in learning theory affects instructional design and continues to inform teaching and learning in university classrooms albeit in less than systematic and comprehensive ways. Lessons from early behaviorism, cognitivism and constructivism have advanced our understanding of learning and informed teaching in all fields. In science education in particular, these lessons resulted in the IPT model that has been applied for over two decades to the study of animation in learning. This research builds upon these lessons by drawing upon techniques used to evaluate learning and learner cognitive process through the IPT lens. Mental model change and procedural data collection, common methodologies in this framework, are used in this study to determine whether and how animation benefits student learning related to atmospheric circulations. Additionally this study seeks to address an important lacuna identified in prior animation research: the determination of the optimal time for learners to examine the dynamic representations depicted in animations.

Prior research has shown that learners given an additional representation of a physical system tend to develop better mental models of the system (Mayer 2001; Lowe 2004). Therefore, the three treatment groups viewing the animations are predicted to outperform control group participants. Moreover, the amount of change between pretest and posttest for the treatment group members should exceed the control group's results. Additionally, prior research has shown that animation viewing is only beneficial to

learners once a sufficient level of prior knowledge has been achieved (Goldman 2003). Therefore, the two treatments groups in which animations are viewed after or while examining the textbook passage should outperform the treatment group viewing the animation before the text. And lastly, Azevedo's work (2007; 2004) examining SRL processes related to learning in multimedia and hypermedia environments predicts that students able to choose when to view the animations should outperform the other treatments groups because the use of metacognitive processes should allow these learners to determine the optimal time to view the dynamic representation. These hypotheses serves as the basis for this study and their evaluation should inform earth science instructors about the use of animations in their instruction of domain novice learners.

Chapter 3: Methodology

Section 1. Target Content Selection and Justification

Prior research has shown that the use of animations and multimedia environments produce little advantages to learner understanding when content delivery is normalized across representational modes (Tversky et al. 2002; Hegarty & Narayanan 1998). However, the variability of published empirical findings suggests that baseline data is necessary across science disciplines to discern if and where possible advantages exist for the dynamic environment (Höffler & Leutner 2007; Russell et al. 1997). Given the ubiquity of the production and inclusion of animations by textbook publishers and their frequent use in science classrooms, this research seeks to determine whether these animations are beneficial to introductory undergraduate science learners and, more importantly, when and how these animations should be employed to increase content understanding. Moreover, this research examines these questions as part of an authentic learning episode embedded within an introductory earth science course at a mid-size Mid-Atlantic Regional Comprehensive University. The content selected for this study was the three-cell model of global atmospheric circulations, which after examining preliminary pilot study results, was subsequently reduced to include only the tropical or Hadley Cell circulation given learning episode time constraints and participants limited prior knowledge of the topic.

Atmospheric circulation was selected as the target content given the macroscopic scale of Hadley Cell circulation, roughly existing between 30° N and 30° S latitude, and the microscopic scale of the atmospheric molecules that constitute the physical matter embedded within the system. Dynamic visualizations of the complex system in the form

of animations appear well-suited as a beneficial representational method to supplement and compliment static diagrams and textual descriptions given the scales involved exceed direct human perception. Moreover, because the circulation system results from differential heating and resulting variances in atmospheric densities directly caused by latitudinal and seasonal differences in incoming solar radiation, additional processes occurring at imperceptible visual and temporal human scales can be visualized by these dynamic representations. The ability to off load working memory function in the coordination and mental simulation of these interrelated processes comprising the system *in motion* is predicted to improve learner understanding by enabling executive control to encode salient features of the atmospheric circulation system neglected while reading textual descriptions and examining static diagrams. Experience teaching this content to first and second year, undergraduate non-science and science majors for over a decade indicates that most students have a difficult time integrating temperature and pressure relationships as well as wind systems functioning across time and space. Students often confuse vertical air motions associated with thermal and dynamic pressure systems and therefore misrepresent resultant weather conditions and horizontal air motion (i.e., wind). Animations enable the learner to view the system repeatedly and dynamically link vertical motions to temperature through the use of colored arrow symbology (e.g., red arrows warm air, blue arrows cold air), to pressure systems through the use of H and L text to represent high and low pressure systems respectively and weather conditions through the use of cloud/precipitation symbology. Resultant horizontal wind directions are shown to move from pressure centers to complete the convective circulations (Figure 1).

Figure 1. Global Circulation Model Including Hadley Cell
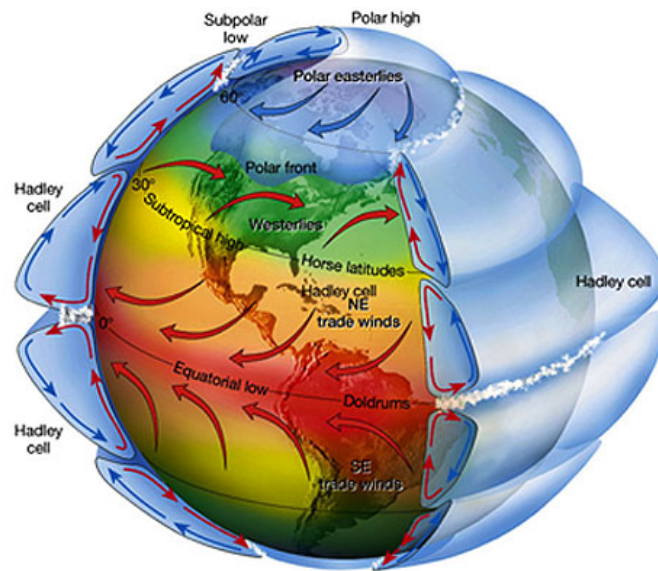
In addition, spatially contiguous depictions of cloud locations with dynamic color-coded and directional arrows should enable learners to directly view and coordinate these inter-relationships with less generative cognitive effort than the mental simulation required between text and static diagrams. Thus, this research attempts to infer the requisite conditions necessary to maximize student understanding when able to view the Hadley Cell and atmospheric pressure processes as depicted in animations.

Section 2. Target Content

The content area selected for student learning in this study initially focused on the general model of global atmospheric circulation for two reasons. First, the model is included in most introductory and advanced earth/atmospheric science textbooks to explain the locations of persistent pressure and wind systems that dictate specific weather occurrences and climatic patterns for the earth (Aguado & Burt 2010; Danielson et al. 2010; Hess & Tassa 2010). Second, the content area was selected due to the system's

common portrayal in textbook publisher animations (Edelson 1996; Lowe 2003). Animations should provide advantages to the learner based on the temporal and spatial scales of the system depicted. Given that most atmospheric constituents are microscopic and invisible at the human scale and that energy and circulation patterns occur at macroscopic scales beyond human perception (without aid of remote sensing devices such as satellites and Doppler radar), the addition of temporal change enables the animation to represent and model dimensions of the target content beyond normal human sensory input. Textual description and static images can and do depict the complex system's processes and resultant patterns, however extended exposition and multiple images and/or diagrams are usually employed to convey the information visible in a single animation sequence. Conversely, animations represent the system in an efficient (i.e., in less than one minute) and integrated way using symbology standard to depictions of weather, temperature and air movement if the learner has been exposed to these symbols through on-air meteorological explanations by their local or a national broadcaster or in the weather section of newspapers. Most animations present general circulation in a single animation with user interactivity enabling the viewer to enact subsystems related to the three cell model: the Hadley Cells, the Ferrel Cells and the Polar Cells. Thus, subsystem content within the animation can be incrementally added by the learner until all of the model's atmospheric motions are observable simultaneously.

The model contained in these animations simplifies some observable atmospheric conditions by removing the impacts of topographic features, migratory pressure systems, land/sea interactions and local geographic settings, however provides provide sufficient explanatory power of the fundamental processes and relationships thereby warranting its

use. Moreover, an understanding of the global circulation model is required prior to the presentation of regional and local modifications to the system, for example monsoonal patterns, the ENSO (El Nino – Southern Oscillation) and rain shadows.

Before presenting the general model of atmospheric circulation to the participants in this study, it was assumed that the learners had a basic understanding of earth-sun relationships, the global radiation balance, atmospheric pressure and wind. These topics and applications of these concepts were presented in prior class meetings and laboratory exercises with the specific focus of explaining atmospheric pressure, the formation and characteristics of high and low pressure systems, why winds occur and the forces that impact wind direction. A detailed description of the pressure and wind content presented prior to the experiment can be found in Appendix A.

Pilot study results indicated that the selected target content contained too much information for content area novices given the time period provided for the learning episode (approximately one hour). Thus, the target content was reduced to a single circulation cell within the model of global atmospheric circulation, the tropical Hadley Cell. A brief description of the Hadley Cell follows.

In the Tropics, or the region defined as lying between 25°N and 25°S in earth-atmosphere science texts, surface winds are recognized as the most consistent system on earth. Surface - or trade - winds, result from the thermally induced convective circulation system called a Hadley cell, one of which operates in each hemisphere between the equator and 25° N and S°. In this case, trade winds converge on the equator where consistently high levels of incoming solar radiation (i.e., insolation). The radiation heats the earth's surface and the overlying atmosphere through conduction and convection (i.e.,

the sensible heat flux) and the latent heat flux associated with prodigious evaporation rates and energy transfer aloft. The resulting high surface temperatures cause lower atmospheric densities as the air expands leading to lower atmospheric pressure. This consistent area of low pressure is termed the intertropical convergence zone (ITCZ). As the trade winds converge and air rises, the air adiabatically cools due to decreasing pressure, releasing enormous quantities of energy aloft as water vapor changes phase and condenses to liquid droplets thereby resulting in abundant cloud cover and precipitation. The surface area beneath the ITCZ is characterized by very warm temperatures, variable winds, and frequent cloud cover. While visible as cloud cover and precipitation on the local scale in discrete locations, the ITCZ is visible in satellite images enshrouding the most of the equatorial globe. Air in the ITCZ rises until it reaches the tropopause where the temperature inversion caused by warmer air the stratosphere effectively precludes uplift and forces air to move north and south latitudinally aloft as the anti-trade winds. While the air is transferred away from the equator, cooling occurs as energy is emitted to space thereby increasing density and atmospheric pressure. The air descends and reaches the surface between 20° and 35° N and S in an area of high pressure called the subtropical high (STH) pressure belt.

Conditions beneath the STH are hot due to adiabatic warming and solar heating, and are clear and dry due to sinking air and the lack of cloud cover and precipitation. Given the long-term implications of atmospheric conditions, deserts correlate strongly with the positions of the subtropical highs (e.g. the Sahara, Mojave, and Australian deserts). The pressure gradient resulting between the ITCZ (i.e. low pressure) and STH (i.e. high pressure) regions initiates and sustains the trade wind circulations in the

Tropics. Air flows from the STH to the ITCZ due to pressure gradient force and is deflected by the Coriolis Effect resulting the northeasterly trade winds in the Northern Hemisphere and the southeasterly trade winds in the Southern Hemisphere. The winds begin relatively dry due to subsidence however obtain significant humidity through evaporation over tropical ocean basins en route to the ITCZ. Aloft, upper level winds, the anti-trade winds, flow in the opposite direction to the surface winds because the pressure system is reversed. As the rising air accumulates above the ITCZ, higher pressures are found while decreased pressure is observed over the STH as air descends. The trade winds, anti-trade winds, ITCZ and STH are all components of the Hadley Cell convective circulation system. Descriptions for the eliminated Ferrel and Polar Cells are found in Appendix B.

Section 3. Participants and Setting

Participants in this study were enrolled in an introductory earth science course at a Mid-Atlantic regional comprehensive university with a student body of approximately 8200 students. The earth science course in which the study was conducted was designed as a general education (i.e., required course) laboratory science course for non-science majors although junior and senior biology majors frequently enroll in the course as an elective for the ecology and marine biology program. The course seated a maximum of 96 students in the lecture hall which met for a total of 150 minutes each week, 75 minutes on two days. The laboratory portion of the class met for 100 minutes one day each week and contained 24 students per section. Thus the lecture course spanned four lab sections with instruction provided by the author in lecture and each laboratory meeting. The

lecture hall and laboratory served as the setting for data collection in this study. A random sample of three students in each laboratory section was selected for process data collection and met the study's author in the same laboratory outside of normal lab time.

<u>*Section 4. Participant Characteristics*</u>

Participants in this study were drawn from two sections of an introductory laboratory science course at a mid-Atlantic regional comprehensive university. While total enrollment in the two sections of the course capped at 192 students, absences during the administration of pre-tests and the treatment conditions in lab meeting reduced the sample size used in the analysis to 175 subjects. Each remaining subject completed a demographic survey prior to the study's pretests to provide background information about participants and to place the sample in the context of overall campus characteristics.

Demographic data indicates that the study's participants are almost evenly divided by gender with male students (e.g., 90 or 51.4%) slightly outnumbering female students (e.g., 85 or 48.6%) unlike the university's student body population where females account for 56.5% of the student body. This difference may be explained by the predominant gender composition of two of the campus' largest majors, education and nursing, which require different lab science coursework in their degree programs. Education majors are required to complete an introductory biology course, an earth-space science course and either an introductory chemistry and physics course while nursing majors must complete biology and chemistry course. Thus, these large and primarily female majors are not enrolled in the course from which this sample was drawn. The

ethnicity of the sample skewed greatly toward individuals who responded white, 91%, on the survey while African-Americans made up the second greatest group at six percent. The remaining three percent identified themselves as Hispanic, Asian or Russian.

Nearly 70% of the participants were aged between 19 and 21 years with 11% reporting an age less than 18 years and 19% reporting an age over 21 years. As the age data indicate, the study sample was weighted more heavily toward juniors (3rd year students) (35%) and sophomores (2nd year students) (31%) than freshmen (1st year students) (12%), with seniors comprising the third largest percentage (22%). While participant age and class standing may be surprising for an introductory course, two possible explanations may explain the sample characteristics. First, many non-science majors delay the completion of their lab science requirement until later in their college career given these courses' perceived rigor. Additionally, the course in which this study was conducted is a requirement for ecology and marine biology majors whose schedule constraints and advising recommendations place the course in either their junior and senior year. Because upper class men and women are able to register for the course first, fewer slots are available to freshmen when their registration period occurs.

Additional demographic, descriptive and attitudinal variables were collected as possible independent variables in the subsequent analyses as well as to provide an expanded background to the participant sample. These variables include student school (e.g., social science, science, business, education) and major (e.g., business administration, psychology, microbiology, etc.), the number of prior mathematics, statistics and laboratory science courses completed, self-reported overall and major grade point averages, whether the participant worked and if so the number of hours worked per

week, attitudes toward prior and current science courses, and whether the student accessed web-based resources including animations for the course.

Science majors comprise 51% of the sample with social science/humanity majors second at 33% followed by business majors at 11%. The remaining students are either social work or physical education majors housed within the education school or students with undecided majors. The science student majority was further revealed by the overall number of science courses taken by study participants. Forty-five percent of the sample reported completing more than three science classes while 33% reported completing one or fewer science courses. However, the pattern of completed mathematics and statistics coursework was notably different. Only 14% reported completing three or more math/stats classes while nearly 62% reported completing one or fewer courses. These measures point toward an apparent science/non-science dichotomy in the overall sample which is explored in the analysis section.

Self reported grade point averages (GPA) indicate 52% of the sample has an overall GPA between 3.0 and 3.99 while 42% report a 2.0 to 2.99 average. Of the remaining students, three percent of the sample reports either a 4.0 or less than 2.0 overall GPA. Grade point average in the subject's major mimics the overall GPA pattern with a slight shift toward higher grades. Nearly 60% of the sample report a "B" major average, 30% report a "C" average, nine percent an "A" average and one percent report a major GPA less than a 2.0. These data provide an interesting baseline from which to contrast the grade expectation reported for the course. Nearly 59% of the sample expected to receive a "B" in the course, more in line with their performance in a majors' course versus their overall coursework while almost 30% expected an "A" greatly exceeding self

reported prior performance in both their major and overall coursework. Only 11%

expected to receive a "C" while no one believed that they would receive a "D" or an "F".

The distribution of grade expectations seems to be supported by the overall

student attitude toward the course in which the study was administered and participants

overall attitude concerning science coursework completed. On a five level Likert-type

scale with one equal to dislike and five equal to like, 63% of respondents selected levels

four and five for the earth science course while only 11% chose levels one and two.

Twenty-six percent reported a neutral level of three. The same pattern was found for all

science courses completed with a slight increase in the dislike values (i.e., levels 1 and 2)

to 15% and a slight decrease in the like values (i.e., levels 4 and 5) to 60%. Overall the

majority of students report a positive attitude toward science classes which is not

surprising given the distribution of reported majors.

Two additional survey questions were included to indicate external, non-academic

commitments that might impact course focus and study habits and thus potentially their

concentration during the experiment. These questions asked whether students were

currently employed and if so the number of hours worked per week. Fifty-three percent

of the sample reported a job with the majority (i.e., 21%) working between 10 and 20

hours per week. Of the remaining working students, 13% reported working less than 10

hours, 11% between 20 and 30 hours while only eight percent reported working greater

than 30 hours per week.

The last two survey questions asked whether students used the online resources

available through the textbook publisher and whether the textbook animations included in

their text had been consulted for prior course content. Nearly 70% reported using

textbook website resources while 55 % reported using the included textbook animations. Thus over half of the participants had chosen to use animations with prior course topics perhaps mitigating the novelty effects on motivation and affect suggested in early animation research (Atkinson et al. 2005). Thus, results discussed below from this study may be seen to have a reduced novelty impact.

Section 5. Experiment Setting

The study was conducted during regularly scheduled lecture and laboratory meetings during a two week period in the second month of the semester. Data collection occurred in both an amphitheater-style lecture hall and earth science laboratory setting. The laboratory is arranged to seat groups of four students around six tables with the instructor located behind a table in the front of the room. Pre-test data were collected in the lecture hall while the learning episode, post-test data and process data were collected in laboratory setting.

The nature of the setting, content material and conceptual understanding expectations of this study coincided with preceding and subsequent course pedagogy and curricula. Students completed the study's tests and learning assignment surrounded by classmates and the researcher in a more authentic setting than the clinical studies more commonly associated with learning from animation research (Stull & Maher 2007; Lowe 2004; Mayer 2002; Brewer 2000). The authentic setting provided in this design differed from common clinical implementations in that students were not recruited from the overall student body to work with the researcher in a one on one setting. Additionally, the target content of the learning episode was selected from the course text therefore students

knew that assessments of their understanding would be administered and impact their course grade (although not the assessments used in this study) initiating intentional rather than incidental learning. The participants read the material and viewed the animations at their own pace given the potential constraint of the laboratory meeting time of one hour and forty minutes (although no participant used the entire time period) therefore enabling the evaluation of their own comprehension prior to receiving the comprehensive post-test assessment tool.

Because the researcher conducting the study was their instructor of record for the course, this study must be considered through the lens of practitioner research including the potential for ethical dilemmas reported in this literature (Fraser 1997). The researcher's bias in this study included the assumption that students would perform at their highest level during the non-graded learning episode, and therefore students not complying with this expectation may be evaluated more harshly in subsequent course assessments. The potential for this bias was reduced by assigning unique ids to the pre-test and post test materials and using these values to link these learning assessments for later analysis. Therefore individual student names were never known or considered during the data evaluation and analysis, insuring anonymity and confidentiality, and reducing the likelihood for future researcher bias. This technique was not possible for students selected for the think aloud data collection given the face-to-face nature of the research setting. However, data transcriptions from this process were assigned unique identifiers to link to the pre-test and posttest materials of these students for subsequent data analysis. Thus as in the preceding case, student anonymity and confidentially existed although the researcher knew the identity of these selected students for the remainder of

the course. Given that these data were used to explore and examine specific cognitive processes employed during the learning episode and not to evaluate the impact of the differing treatment conditions, the research design itself attempted to reduce this bias.

Section 6. Research Design

This study used a pretest-posttest quasi-experimental design to measure conceptual understanding and conceptual change given the strong level of internal validity and ability to control for confounding variables. Quantitatively, a mixed factorial research design was used to statistically evaluate within and between group pre-test and post-test differences to ensure randomization between groups and reveal causal effects resulting from the different treatment conditions. Although this design may produce questions of external validity, the seven day period between the pre-test and post-test assessments should minimize any possible priming or interaction (Shuttleworth 2009). Moreover, this design does not introduce a possible instrumentation problem due to the change in assessment tool and allows for the ability to calculate an effect size based on change from pre-test to post-test. Planned contrast ANOVAs were performed to compare group mean performance differences based upon treatment condition as well as to compare the impact of the treatment on groups of learners categorized by type determined by pre-test assessments. After analyzing mean groups differences, a regression analysis was used to control for individual differences discovered in the pre-tests and to predict the amount of variance in the posttest performance resulting from the treatment effect and any additional, significant predictor variables.

A secondary analysis was also conducted for a twelve student sample to determine whether differences in if self-regulated learning (SRL) behaviors occurred by learners during the learning episode and whether these processes varied by treatments condition. These data were collected using think aloud protocols given the techniques' frequent and successful use to reveal cognitive processes during learning (Ericsson 2006; Azevedo et al. 2005). Coding followed the scheme described in Azevedo and Cromley (2004) (Appendix D) and given the small sample size, these data were compared qualitatively to discern any recognizable patterns.

Section 7. Materials

Testing materials used in this project included a short demographic questionnaire and a pre-test and post-test developed by the author who holds a graduate degree in physical geography with a specialization in climatology. The demographic questionnaire was developed to determine the overall characteristics of the student sample and to place the sample into the broader context of the University's student body. Questions pertaining to student age, gender, ethnicity, academic year, major, grade point average, prior math and science coursework, work hours if any, perceptions of prior and current science coursework, and textbook supplement use were obtained via this instrument (Appendix C). These measures were evaluated as potential explanatory variables in statistical analyses conducted in the results section.

Section 8. Assessments

The participant demographic questionnaire, verbal abilities and spatial abilities, and target content multiple choice, essay and diagrammatic pre-tests were administered in the lecture hall one week prior to laboratory meetings. Student conceptual understanding was evaluated according to relative understanding of the pressure and wind systems depicted by the animations and with their associated meteorological and climatic conditions. Two assessments types were employed before and after the learning episode. The first assessment examined performance change on the same pre- and posttests of declarative and procedural knowledge using twenty multiple choice questions, between the control group and the three experimental groups. The second assessment examined pre-test and post-test essay and diagrammatic explanations of the target content for categorization into mental models levels. Resulting mental model and level changes following the intervention for students in each group were compared for qualitative differences in an effort to confirm and support understanding differences noted in the statistical evaluation. Additionally, a sample of three students per experimental group, treatment and control (n=12) was selected for the collection of "think alouds" to illuminate SRL processing differences between groups. These students completed the same experimental sequence of pre-test and posttests with the additional of a digital audio recorder that recorded their vocalizations during the learning portion of the exercise.

Section 9. Content Knowledge Assessment

The identical pretest and posttest assessment instrument was designed to comprehensively evaluate students' understanding and knowledge of the target model of

Hadley Cell circulation (Appendix D). Unlike the instruments employed in most animation research that contain only multiple choice or short answer questions, this test included twenty multiple choice questions, four open-ended term and concept identification questions, an essay question and two diagram questions designed to evaluate students' declarative and procedural and inferential knowledge of the content material prior to and following the experimental intervention. The comprehensive nature of the assessment tool is relatively unique and enables the participant to externally represent target content understanding across many forms thus allowing the researcher to evaluate the depth of the learners' understanding. Moreover, the assessment test was designed in a manner similar to exam structures administered across geosciences courses and university campuses, therefore the participants should theoretically be familiar with the format and tailor their comprehension strategies to successfully navigate the assessment tool. The multiple choice and term identification questions were designed to evaluate declarative knowledge in the sample while the essay and diagram questions require the participant to link declarative content with causal chains implying procedural understanding of the interrelationships among the target concepts. An example of a multiple question and term identification question designed to assess declarative knowledge follows.

> Global atmospheric circulation is driven by:
>   a) latitudinal energy imbalances
>   b) spring and neap tides
>   c) oceanic circulations
>   d) earth's distance from the sun
>   e) the moon's gravitational attraction
>
> **Short Answer:** Define and describe each of the atmospheric features listed below.

a) Intertropical Convergence Zone

Procedural understandings were assessed by the following essay and diagram question.

>**Essay:** Describe the generalized model of Hadley Cell as discussed in your learning materials including pressure and wind features. Within your discussion, identify the locations of the major pressure centers and the dynamic and thermal mechanisms associated with their locations. Also describe the impact of these pressure systems on expected weather conditions as understood by vertical atmospheric motions. Include in the discussion the seasonal changes expected in this pattern over the course of one year. Use diagrams to describe the processes exemplified in the essay.

>**Diagram:** Draw the generalized pattern of the Hadley Cell on the first globe below. Label the pressure centers, wind directions, and average weather conditions.

Globe 1: General pattern (reduced)
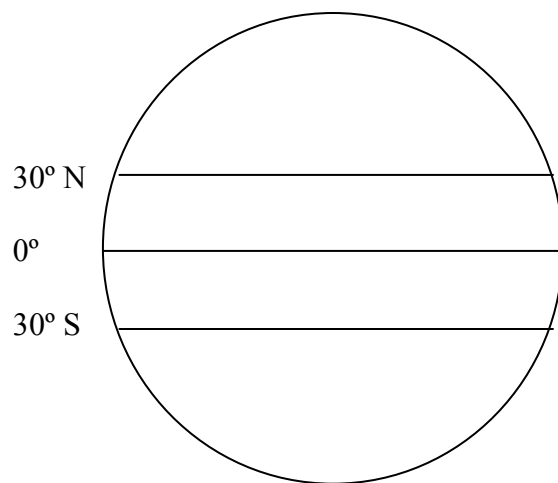
The final diagram question assessed the participants' ability to transfer the Hadley Cell pressure and wind pattern depicted in the learning materials in the equinox position to a different point in earth's orbit not discussed or represented in the copied text. Thus the learner was required to coordinate both the text and diagrammatic representations of

the system and apply their understanding of the driving mechanisms to a future state and therefore a different position on the globe. The transfer diagram could not be completed correctly through simple memorization but required the participant to infer circulation system movement based on changes in the causal mechanism solar declination and incoming solar radiation receipt. Moreover, the participant would need to realize that the vertical and horizontal air motions associated with the pressure systems would remain constant even with the latitudinal shift. A correct response on the transfer question should theoretically indicate a comprehensive understanding of the target content and therefore predict high scores on both posttest measures.

The multiple choice questions were scored on a scale of one to twenty points based on the total number of correct responses for each question and were evaluated separately from the open-ended and diagrammatic responses. Open-ended responses were cumulatively evaluated to determine the students' expressed mental model of the target content on an ordinal scale ranging from zero (e.g., completely blank and/or erroneous responses) to seven (e.g., complete system understanding) derived from the range of observed student responses. The mental model assessment technique was selected given its widespread usage by cognitive psychologists and science education researchers to evaluate cumulative understanding of complex systems and its ability to comprehensively integrate multiple external representations generated by learners in experimental conditions into an ordinal rank for evaluation and comparison (Azevedo et al. 2004; Greene & Azevedo 2007; Libarkin et al. 2003). Table 2 below describes the components of each model level defined in this study.

Table 2. Experiment Derived Mental Model Levels

Level	Component Model Description
I. No understanding of generalized atmospheric circulation (Blank or nonsensical)	Most wind directions and pressure systems omitted or incorrectly represented by location and characteristics
II. Weak understanding of generalized atmospheric circulation (incomplete and inaccurate)	Few wind directions and pressure systems correctly represented by location and characteristics
III. Moderate understanding of generalized atmospheric circulation (More incomplete than inaccurate)	Some wind directions and pressure systems correctly represented by location and characteristics with prominent omissions
IV. Strong understanding of generalized atmospheric circulation (Mostly complete with few inaccuracies)	Most wind directions and pressure systems correctly represented by location and characteristics with few omissions
V. Near complete understanding of generalized atmospheric circulation (Complete with few inaccuracies)	All wind directions and pressure systems correctly represented by location and characteristics with few omissions
VI. Complete understanding of generalized atmospheric circulation (Very few omissions and inaccuracies)	All wind directions and pressure systems correctly represented by location and characteristics with very few omissions
VII. Complete understanding of generalized atmospheric circulation (No omissions and inaccuracies)	All wind directions and pressure systems correctly represented by location and characteristics

Learning materials assembled for the experiment were composed of photocopies of text and diagrams selected from the course textbook describing air pressure and Hadley Cell circulation and textbook publisher constructed animations representing the same content (Hess & Tassa 2010). The photocopied materials contained nine textbook pages of approximately 3000 words and 19 static diagrams portraying the target content and the textbook author's presentation was evaluated to be similar to explanations provided in most introductory earth science textbooks. The length of the required text and number of diagrams employed in this study is uncommon in most science learning research especially prior work in animation use. However, in the context of an introductory undergraduate science course, the material is consistent with curriculum design and content expectations given the period of time devoted to the experiment in the

course. These courses require extensive and frequent textbook reading, especially in courses in which students' exhibit limited to non-existent prior knowledge as was the case for the content included in this study.

The selection of the target content sought to align the research study to real world learning expectations in a university environment thus increasing ecological validity in contrast to most prior animation research where subjects are selected from the general student body or from specific major tracks (e.g., education and psychology) and are asked to watch a two to five minute animation about a system outside of their knowledge base. In most circumstances, these animations contain limited text, verbal and/or static descriptions of the depicted system and the focus of the researcher is to increase sample size given the short period of learning episode at the expense of content depth. Similarly, posttests in these studies tend to shy away from mental model assessments given the paucity of potential data types collected following the more brief interventions. This study's design seeks to discern more typical learning processes employed by undergraduates in the context of a real introductory science course and evaluate the understanding outcomes attributed to the inclusion of an additional representation mode. Rather than simply assess learning from one or perhaps two mediums, this study incorporates three representational modes most prevalent in content delivery forms accessed by college learners outside of direct instruction.

Two textbook supplied animations were selected for the treatment groups in this study. The first animation displayed horizontal and vertical circulations associated with Northern Hemisphere low (i.e., cyclone) and high (i.e., anticyclone) pressure circulations. The individual animations for each circulation type were viewed when users depressed a radio button and a narration synchronized to element movements accompanied the dynamic representation. The user also had the ability to remove the animation labels by depressing another radio button. Each animation subsystem, cyclone and anticyclone, could be viewed in its entirety in 10 seconds, 20 seconds total, although the default design allowed the animation to continue to play until users de-selected a buttons or closed the program. The second animation depicted the general model of global circulation with design and controls constructed and implemented in a similar fashion. Three buttons, each linked to either the tropical cell (i.e., Hadley) cell, the midlatitude cell (i.e., Ferrel) or the Polar cell, controlled the onset of each animation subsystem. However to view all subsystems simultaneously, the user had to depress the buttons in sequence top to bottom. The first active button was the Tropical circulation cell and this subsystem was completed after five seconds. The midlatitude circulation button was activated next with the full depiction viewable in three seconds. Lastly the Polar system was clickable and it too could be viewed in three seconds. A fourth button was available after all subsystems were displayed depicting the subtropical and midlatitude jet streams and it was also viewable in three seconds. The animation design also included buttons to remove feature labels while maintaining dynamic movements and a button to clear animations labels if desired. This button was available for use throughout any subsystem

sequence. Thus, the entire circulation system could be viewed in less than 20 seconds if the learner depressed each button upon completion of each subsystem and like the first animation, played continuously until users turned off elements or closed the program.

Section 11. Materials: Measures of Verbal Ability and Mental Rotation

Pre-tests evaluating verbal ability and the spatial ability mental rotation were obtained and administered due to prior research identifying these factors as predictors of science comprehension and achievement especially relevant to complex system and animation understanding (Hegarty& Kozhevnikov 1999; Holliday et al. 1984). The verbal ability test was a 48 question standard vocabulary test (V-3) from the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al. 1976) selected given its strong record of correlation with assessments of understanding and prediction for content comprehension when reading science texts (Holliday, Brunner, & Donais 1977). The spatial abilities test selected was the re-drawn Vandenberg and Kuse Mental Rotation Test (Peters et. al. 1995). Copies of V-3 test are available from Educational Testing Service while the re-drawn Vandenberg and Kuse Mental Rotation Test is available from the primary author (Peters et. al. 1995).

Prior research has shown that the ability to mentally manipulate and transform objects is related to the ability to mentally simulate static diagrams and to identify important components in a system's causal chain (Hegarty & Kozhevnikov 1999). While three specific spatial abilities, visualization, rotation, and orientation have been shown to be important in mental animation, visualization abilities measured by the Paper Folding Test VZ-2, spatial orientation abilities measured by the Guilford-Zimmerman Spatial

Orientation Test and mental rotation abilities measured by the Vandenberg's Mental Rotation Test, only mental rotations were used in this study (Hegarty & Kozhevnikov 1999). The author chose this particular spatial ability given the nature of the animation designs viewed by experiment participants. Arrows symbolizing vertical and horizontal motions were initiated from the earth's surface, rose and then moved north and south of the equator before sinking to the surface and returning to their original position. Circulation in each hemisphere's Hadley Cell and implied in the text and static diagrams was a key concept for novice learners in the experiment. Therefore the ability to simulate rotation through mental animation was hypothesized to enable deeper conceptual understanding from each representation type resulting in better performance on posttest assessment measures. Both animation designs did not allow users to move the perspective point from which the animations were viewed above the surface of the globe thereby changing wind arrow directionality or north-south orientation. Therefore the abilities measured in the spatial orientation (or perspective taking test) would not provide information useful to predict relevant information extraction from the animations used in the treatment. Similarly, the paper folding test assesses the learners' ability to visualize transformed planar objects. Animations used in the study portrayed the earth as a three-dimensional object with an invariant frame of reference. Thus the surface of the earth on which the target concepts were depicted did necessitate the learner to transform objects into planar representations.

Section 12. Scoring Mental Model Levels

Assessments of mental models can be difficult given the inability to directly examine the learner's cognitive structure necessitating the production of an external representation of the target content for evaluation (Norman, 1991). The pretest and posttest assessment tool used in this experiment required participants to represent their understanding of the target content though both textual and diagrammatic descriptions. These novice representations were evaluated against the content model presented in the provided text section, considered to be the expert model for this study (Vosniadou & Brewer 1992). As prior researchers have noted (Coll and Treagust 2003; Norman 1991) when comparing models to models, interpretations and assessments are difficult given the" messy, sloppy, incomplete, and indistinct structures that people (i.e., novice learners) have" (Norman 1991; p. 14). Moreover, researcher interpretations of these models are mediated by their ontological and epistemological lens therefore independent evaluation improves reported results (Coll & Treagust 2002). By operationally defining seven mental model levels (Table 2) based on the range of student representations observed, the author rated each participants' externally represented model from both the pretest and posttest an ordinal scale between zero and seven. An independent rater with a graduate degree in the atmospheric science examined one-half of the pretests and posttests, randomly selected from each of the four conditions, and rated the responses on the same ordinal scale. Both raters were unaware of the treatment condition when evaluating the mental models. Rater scores agreed in approximately 93% of the cases and disagreements were resolved on a case by case basis through discussion and score revision to the agreed upon level.

Section 13. Experimental Procedure

The experiment was conducted in two university lecture sections composed of 96 students each randomly assigned by course registration to one of four experimental groups: the control group – no animations and three treatment groups – animations before the text and diagram packet, animations after the text and diagram packet, and user determined animation examination. Each lecture 100 minute section was divided into four 24 person laboratory sections and laboratory registration served as random assignment to an experimental or control condition.

The demographic test, verbal test and spatial abilities tests (completed in five minutes, ten and ten minutes respectively) were administered during the lecture class one week prior to the laboratory experiments utilizing the instruction sets provided with each test. Five total minutes were used to read directions for each of the pretests, with the exception of the content pretest, and to allow all students be seated prior to the data collection. The content pre-test was completed in the remaining forty-five minutes of the lecture meeting. The script used to collect the content pre-test data stated "This research is being conducted by Dr. William Holliday at the University of Maryland, College Park and Daniel Harris at Salisbury University. We are inviting you to participate in this research project because you are a student in an introductory geosciences course. The purpose of this research project is to determine whether animation viewership and/or the absence of animation viewership improve understanding of atmospheric circulation systems. If you decide to participate in this experiment, please complete the following prior knowledge assessment to the best of your ability given your current understanding

of the topic. If questions arise, raise your hand and I (i.e., the instructor) will walk to you and answer your question."

During the learning episode administered in the lab, students received five minutes of instruction on the purpose of the research (a review of the directions delivered the preceding week during class), the goal for examining the learning materials and how to employ the animations on the laptops if assigned to an experimental group during the learning episode. The instructions stated "This research is being conducted by Dr. William Holliday at the University of Maryland, College Park and Daniel Harris at Salisbury University. We are inviting you to participate in this research project because you are a student in an introductory geosciences course. The purpose of this research project is to determine whether animation viewership and/or the absence of animation viewership improve understanding of atmospheric circulation systems. If you decide to participate in this experiment, please accept the packet of learning materials copied from your course textbook. When instructed, open the packet and read the text and examine the diagrams carefully for content understanding. Some of you will have access to a laptop containing animations depicting the target content at specific times during the learning period. These animations are deployed by double-clicking on the icons on the desktop and depressing the buttons embedded in the animations. If questions arise, raise your hand and I (i.e., the instructor) will walk to you and answer your question. Once you feel comfortable with your understanding of the content, turn off the laptop, return the text and diagram packet to the instructor and you will receive a post test assessing your understanding of Hadley Cell circulation."

Following the receipt of the instructions, each participant in each group received a packet containing text and diagrams describing and depicting Hadley Cell circulation. The instructions stated that this study is interested in examining how students learn introductory earth science using multiple representations, text and diagrams (and animations for the experimental groups). Students were told to read the text for understanding and examine the diagrams provided to learn as much as possible about the idealized pattern of Hadley Cell circulation during the laboratory session. Students selected for the think aloud protocol were e-mailed an alternative meeting time to complete the experiment in the laboratory setting and provided with an additional set of instructions related to the think aloud protocol. If animations were used in their treatment condition, instructions were included to deploy the animations on the provided laptop. Members of the user determined animation group (i.e., the self regulated learning group) received the following set of instructions. "Read the textbook materials provided, including the reproduced diagrams, to learn as much as you can about the nature of wind and the Hadley Cells in global atmospheric circulation. The laptop in front of you has two sets of animations playable through the internet browser which is already operational and accessible on the toolbar. Just click on the icon and press the buttons on the animations to display specific circulations. You can turn off the animations with the clear all animations button. *You may look at the animations at any time while you read the text packet.* If you have any problems and/or questions concerning the animations, please ask me. While you read and examine the graphics tell me what you are thinking. If you are silent for over 45 seconds, I will ask you to tell me what you're thinking. When you feel that you understand the content, let me know and we'll move on to the next phase of our

experiment." The phrase in italics was modified for the animation before reading and the animation after reading groups to define the appropriate time to examine the animations.

Participants' self evaluated their understanding of the target concepts during the learning episode and once they were comfortable with their understanding, received the posttest assessment tool. Students completed the posttest which was identical to the content area pre-test administered a week earlier with the addition of a diagram for students to predict Hadley Cell patterns on the June Solstice or system shift to the Northern Hemisphere. Students completed the posttest in the remaining laboratory time and in every case, the learning episode and tests were completed within the given time period without anxiety.

Section 14. Experimental Groups and Expected Results

Three possible timing sequences for animation viewership are defined in this research project. In the first test sequence, the animations are viewed prior to the examination of the text and static images explaining the Hadley Cell circulations. The group was given between five and ten minutes based on user preference to examine the animations, multiple times if desired, before proceeding to the packet of text and diagrams. In this group, the animations should theoretically serve to activate any prior knowledge of the systems' behavior and functioning and/or should prompt learners to seek explanations in the text packet for events or actions within the animations that were not understood. Given the lack of prior knowledge concerning Hadley Cell circulation, animation viewing at this time is hypothesized to be detrimental to understanding given the complexity and interrelationships of the systems components and the lack of top

down guidance to selectively encode relevant information. Learners in this condition are expected to perform at the same level as the control group and at a lower level compared to the two remaining experimental conditions.

In the second treatment group, the animations are available for viewing after the text and diagram materials have been examined. Thus the depictions in the animations are hypothesized to be more easily understood given the greater likelihood of prior knowledge guiding the selective encoding of relevant components and interrelationships. Moreover, the composite Hadley Cell animation may theoretically serve as an organizational structure for the construction of the target content mental model. Each declarative element encoded during the text and static image examination should theoretically be ordered and organized by the animation sequences and linked to each other by the dynamic motions depicted between component parts. This group's posttest performance is expected to exceed both the control group and animation first group.

In the third treatment condition, students were able to view the animations at any time during the learning episode. Thus, these students should theoretically be able to coordinate each representational type (e.g., text, static diagram, animation) concurrently if so desired, clarifying any potential conceptual misunderstandings between representational forms. Moreover, knowledge construction should theoretically proceed in a more efficient manner given that learners can employ metacognitive monitoring processes to compare and contrast explanations across representational forms. Thus evaluations of task demands and monitoring understanding should result in strategy use coordinating all representational forms to reach the learning goal. Moreover, if animation complexity overwhelms working memory capacity, the learner is able to re-visit passages

and diagrams to deconstruct dynamic relationships illuminated in the animation to component parts. It is predicted, based on prior research in self regulated learning and multimedia learning and design emphasizing interactivity and user control, that learners in the third condition will achieve greater understanding of the target content due to their ability to coordinate each representation type while metacognitively monitoring their learning

Chapter 4: Results

Section 1. Pilot Study Results and Modifications

A preliminary study, excluding process data collection, was conducted in the semester prior to the data collected for this study to evaluate implementation logistics and the assessment tools developed to evaluate student understanding. Pilot study administration followed the same sequence described in the this study's procedural implementation with pre-test administered in the lecture meeting one week prior to the laboratory delivered treatment. Student comments following the laboratory learning episode and assessment results indicated that the three-cell model of global circulation was too much content for the participants to apprehend given the time constraints of the lab meeting period. Mean scores on the twenty question multiple-choice posttest were less than one point higher than pretest results across treatment groups (e.g., pretest $M =$ 7.6 vs. posttest $M = 8.3$). Mental model gains between the pretest and posttest were similarly negligible with a difference of less than one level (e.g., pretest $M = 0.9$ vs. posttest $M = 1.4$). Moreover, 75% of posttest assessments received a mental model score of two or less because significant portions of the assessment tool, especially the diagram questions were left blank. These results indicated that novice participants were not able to encode the target content given the complexity of the system and the length of time allowed to read the text and diagram passage for comprehension. Therefore, model content was significantly reduced to include only the tropical Hadley Cells from the three cell model of global circulation.

Results are presented and discussed in the following two sections based on the specific research questions addressed and the assessment tools employed to measure treatment effects. The first section contains the quantitative analyses used to evaluate the two primary research questions examined in this study: 1) Does animation viewing, in supplement to textbook explanations and diagrammatic representations, increase target content understanding? And 2) Does the timing of animation viewing in relation to examining the textbook materials affect resultant content understanding? These analyses utilized parametric statistics based on the underlying distribution of the variables and the level of measurement employed in assessing participant understanding. The pre-test measures used across all analyses included the V-3 vocabulary test to infer science text comprehension, Vandenberg and Kuse's Mental Rotation Test to evaluate participants' ability to identify three-dimensional features from differing perspectives, a multiple choice assessment to measure participants' prior knowledge and an open-ended, comprehensive assessment of prior knowledge which included short answer questions, essays, a diagram and a transfer diagram. Posttest measures to evaluate understanding changes following the intervention included the results from the same multiple choice and mental model assessments as well as understanding gain variables constructed by calculating the differences between posttest and pretest scores for each participant on the multiple choice and mental model assessments.

The second section of the results presents a qualitative analysis of process data collected during the think alouds for the sub-sample of twelve students. Three students in each treatment condition and the control group were randomly selected for audio recording during the learning episode. The audio recordings were transcribed to text files

and coded based on Azevedo et al.'s (2004) Self Regulated Learning behavior framework to examine whether differences in cognitive processes could by identified between the treatment and control groups.

Azevedo's framework partitions Self Regulated Learning behaviors into five classes, planning, monitoring, strategy use, task difficulty and demand and interest. Within each of these classes, specific variables are identified and described based on empirical observations of learners. Planning behaviors are utilized at the onset of the learning episode to identify and select goals, retrieve prior knowledge relevant to the problem or learning goal, and to coordinate cognitive operations in a hierarchal fashion or partition the goals into discrete steps toward the desired learning outcome. Monitoring behaviors are employed to evaluate information acquisition and understanding in relation to the goals and sub-goals developed in the planning process. Information acquisition to achieve planned goals occurs through an array of strategies employed during the learning event and in some cases specific to the information representations accessed. Strategy use also includes behaviors employed in any learning environment including summarization, note-taking and memorization as well as higher order cognitive processes whereby presented content is used to construct inferences, hypotheses and to guide knowledge elaboration specific to planned learning outcomes. As learning is undertaken, task difficulty and demand are assessed to strategically allocate resources, such as time spent with specific content representations, and to evaluate the information in the content representation to determine whether it adequately meets the current learning goal. Azevedo's last class, interest, simply evaluates whether the learner verbally expresses interest in the content when navigating the learning episode.

Section 3. Quantitative Results: Hypothesis Testing

To facilitate instrument comparisons and statistical model interpretations, pretest and posttest scores utilized in the quantitative analyses were transformed to percentages. Table 3 contains the descriptive statistics, including the mean, standard deviation and range, for each variable in this format.

Table 3. Descriptive Statistics

Variable	Definition	Mean	Std. Dev.	Min	Max
Student Ability					
V-3 score	V-3 Kit of Factor-Referenced Cognitive Tests (%)	44.94	10.87	16.67	79.17
MR score	Vandenberg and Kuse Mental Rotation Test (%)	52.14	23.25	4.00	100.0
MM prescore	mental model assessment (%)	9.22	17.53	0.00	85.71
MC prescore	multiple choice assessment (%)	37.91	16.83	5.00	80.00
Dependent Variables – Post Treatment Scores and Differences					
MM postscore	mental model assessment (%)	34.12	26.12	0.00	100.0
MC postscore	multiple choice assessment (%)	50.00	20.21	10.00	95.00
MC gain	multiple choice posttest minus multiple choice pretest	12.09	12.27	-15.0	45.00
MM gain	mental model posttest minus mental model pretest	24.90	22.55	-14.3	71.43
Student Demographics					
Gender	Female = 1, Male = 2	1.51	0.50	1.00	2.00
Race	White = 1, African-America = 2, Hispanic = 3, Asian = 4, Russian = 5	1.14	0.57	1.00	5.00
Age	$\leq 18 = 1, 19 = 2, 20 = 3, 21 = 4, \geq 21 = 5$	3.10	1.29	1.00	5.00
Year	Freshmen = 1, Sophomore = 2, Junior = 3, Senior = 4	2.66	0.96	1.00	5.00
School	Business = 1, Liberal Arts = 2, Science = 3, Education = 4	2.49	0.76	1.00	4.00
Science major	Science major = 1, other = 2	1.49	0.50	1.00	2.00
GPA overall	< 2.0 = 1, 2.0 to 2.99 = 2, 3.0 to 3.99 = 3, 4.0 = 4	2.55	0.61	1.00	4.00
GPA major	< 2.0 = 1, 2.0 to 2.99 = 2, 3.0 to 3.99 = 3, 4.0 = 4	2.77	0.62	1.00	4.00
Science course	Number of science courses completed	0.49	0.50	0.00	1.00
Math course	Number of mathematics and/or statistics course completed	2.47	1.36	1.00	5.00

	Yes = 1, No = 2	1.47	0.50	1.00	2.00
Job	Yes = 1, No = 2	1.47	0.50	1.00	2.00
Work hours	None = 1, < 10 = 2, 10 to 20 = 3, 20 to 30 = 4, > 30 = 5	2.20	1.35	1.00	5.00
Course attitude	Likert-type scale 1 (dislike) to 5 (like)	3.72	0.96	1.00	5.00
Science course attitude	Likert-type scale 1 (dislike) to 5 (like)	3.60	0.97	1.00	5.00
Grade expected	A = 1, B = 2, C = 3, D = 4, F = 5	1.82	0.62	1.00	3.00
Internet use	Textbook web site use: Yes = 1, No = 2	1.31	0.46	1.00	2.00
Animation use	Textbook animations: Yes = 1, No = 2	1.45	0.50	1.00	2.00

Because the focus of the first research question was to evaluate differences in student understanding following differing instructional interventions, the descriptive statistics presented in Table 2 disaggregate the pretest and posttest scores presented in Table 4 by treatment group to provide a general idea of the distribution and characteristics of these data.

Table 4. Descriptive Statistics by Treatment Group

	Control		Treatment 1		Treatment 2		Treatment 3	
	Mean	Stand. Dev.	Mean	Stand. Dev.	Mean	Stand. Dev.	Mean	Stand. Dev.
Student Ability								
V-3 score	43.12	11.81	44.86	10.33	45.59	9.67	46.16	11.67
MR score	48.93	23.06	48.18	23.01	55.26	24.00	56.27	22.57
MM prescore	11.30	22.17	5.40	13.03	9.30	20.65	11.04	12.27
MC prescore	38.14	15.89	36.22	17.49	38.02	17.43	39.32	16.90
Dependent Variables – Post Treatment Scores and Differences								
MM postscore	27.24	24.31	23.81	21.96	37.87	24.72	47.73	27.03
MC postscore	45.47	18.45	46.00	17.76	48.02	21.93	60.45	19.37
MC gain	7.33	9.15	9.78	12.34	10.00	13.23	21.14	9.20
MM gain	15.95	14.01	18.41	20.34	28.57	21.37	36.69	26.75
Student Demographics								
Age	2.91	1.29	3.16	1.30	2.86	1.26	3.48	1.25
Gender	1.60	0.49	1.44	0.50	1.37	0.49	1.63	0.49
Race	1.21	0.77	1.18	0.39	1.00	0.00	1.18	0.76
Year	2.48	0.97	2.86	0.90	2.51	1.08	2.80	0.85
School	2.51	0.81	2.26	0.82	2.50	0.68	2.70	0.67
Science major	1.51	0.51	1.65	0.48	1.48	0.51	1.32	0.47

GPA overall	2.55	0.68	2.49	0.55	2.74	0.58	2.41	0.58
GPA major	2.66	0.69	2.82	0.62	2.95	0.65	2.65	0.48
Science course	3.12	1.69	3.29	1.60	3.49	1.56	4.09	1.31
Math course	2.44	1.28	2.31	1.43	2.29	1.33	2.82	1.39
Job	1.49	0.51	1.36	0.48	1.58	0.50	1.48	0.51
Work hours	2.05	1.23	2.58	1.50	1.95	1.25	2.20	1.36
Course attitude	3.86	0.83	3.87	0.89	3.74	0.93	3.41	1.11
Science course attitude	3.56	0.96	3.51	1.04	3.74	0.95	3.59	0.92
Grade expected	1.77	0.61	1.89	0.53	1.67	0.64	1.93	0.66
Internet use	1.38	0.49	1.27	0.45	1.36	0.49	1.23	0.42
Animation use	1.41	0.50	1.41	0.50	1.45	0.51	1.50	0.51

Pretest scores by treatment group were evaluated for statistical differences by one-way ANOVA because the quasi-experimental design of this study did not allow for the random assignment of participants to the control and treatment groups. Levene's Test of the Homogeneity of Variances indicated that this assumption was met therefore enabling the use of ANOVA analysis. In all four pretest cases, the V-3 Standard Vocabulary Test (i.e., $F(3, 171) = .636$, $p = .593$), Vandenberg and Kuse's Mental Rotation Test (i.e., $F(3, 171) = 1.431$, $p = .236$), the prior knowledge multiple choice assessment (i.e., $F(3, 171) = .253$, $p = .859$) and the comprehensive mental models assessment (i.e., $F(3, 171) = 1.074$, $p = .362$), no statistically significant differences in group means were discovered prior to the treatment. These results indicate that participant characteristics as measured by verbal ability, mental rotations and two measures of prior knowledge were found to be consistent across each laboratory section prior to the intervention thereby enabling for an analysis of group differences on posttest instruments following the intervention.

Given the study's first research question, an ANOVA was conducted to determine whether multiple choice and mental model posttest differences existed between four

groups in the experiment. The omnibus ANOVA test for the multiple choice posttest indicated that at least one group's mean differs, $F(3, 171) = 5.814$, $p = .001$, as did the test between mental model posttest scores, $F(3, 171) = .8.607$, $p = .000$. Since significant differences on both posttest measures were found, *a priori* planned comparisons were conducted to identify specific group mean differences. The first Multiple Comparison Procedure (MCP) selected was Dunnett's Comparison Method to evaluate the special pair-wise comparisons between the control group and each treatment group. Dunnett's Method provides an adequate level of control for family-wise error rates and enables for directional hypothesis testing, important given the research hypothesis that animation viewing will increase understanding. Results indicated that only treatment group three, the self regulated learning treatment group, differed significantly from the control group's mean on the multiple choice posttest, ($M = 14.99$, $SD = 4.17$), $F(3, 171) = 5.814$, $p = .001$. The same planned comparison method was used to examine group differences on the more comprehensive mental model posttest assessment. Dunnett's method again yielded a significant difference between the control group and treatment group three, ($M = 20.5$, $SD = 5.3$), $F(3, 171) = 8.607$, $p = .000$ but found no difference with group one and two. However, with a significance level of .058, the difference between the control groups and treatment group two, animation viewing after reading was very nearly significant at the 0.05 level.

After finding that the self regulated learning treatment groups' posttest mean scores were significantly higher than the control groups' scores, both learning gain variables were compared by treatment group with the control group. Using the multiple choice posttest gain, differences were only found between the control group and the self

regulated learning treatment group, (*M* = 13.8, *SD* = 2.4), $F(3, 171) = 13.452$, *p* = .000.

However, when using the mental model gain variable, a significant difference was also

found between the control group and treatment group two, the animation after group, (*M*

= 12.6, *SD* = 4.6) in addition to treatment group three, (*M* = 20.7, *SD* = 4.5), $F(3, 171) =$

8.976, *p* = .000.

Results from the initial analysis indicated that across both the posttest and

learning gain measures, the self regulated learning group, on average, scored more highly

than the control group in which no animations supplemented the text and diagram packet.

And in the case of the mental model posttest, treatment group two, animation after

reading, showed a significant difference in learning gains as compared with the control

group and very nearly a difference on the overall posttest score. These findings indicate

that participant performances in the second and third treatment group varied from the first

treatment group, animations prior to reading, as well as the control group.

These results indicate that animations, when viewed by the learner following the

acquisition of sufficient prior knowledge, increase student understanding of Hadley Cell

circulation. To evaluate performance differences between individual treatment groups,

the subject of this study's second research question, a second set of pair-wise planned

comparisons was conducted.

Given that the comparisons of interest were planned, non-orthogonal directional

and included pair-wise instances, the Dunn-Sidak method was selected to compare the

treatment groups' mean posttest and learning gain scores. To control for the family-wise

error rate and to minimize the likelihood of a type I error, alpha was set to 0.017 for each

of the three planned comparisons (.05/3). The first comparisons examined the differences

on both posttest measures. Treatment group three's multiple choice posttest scores were significantly higher than treatment group one's scores, (M = 14.5, SD = 4.1), $F(3, 171)$ = 5.814, p = .001 however treatment group two's scores were not significantly higher than treatment group one's scores. The same result was discovered using the mental model posttest. Treatment group three's mean was significantly higher than treatment group one's mean, (M = 23.9, SD = 5.2), $F(3, 171)$ = 8.607, p = .000.

Results using the multiple choice learning gain measure found a significant difference between group three's mean performance and both treatment group one and two. Treatment group three scored higher than treatment group one, (M = 11.4, SD = 2.4), $F(3, 171)$ = 13.452, p = .000 and treatment group two, (M = 11.1, SD = 2.4), $F(3, 171)$ = 13.452, p = .000. However, using the mental model learning gain measure, treatment group three's mean was only significantly different than treatment group one's mean, (M =18.3, SD = 4.5), $F(3, 171)$ = 8.976, p = .000. The figures below graphically illustrate the group differences on both posttest and learning gain measures.

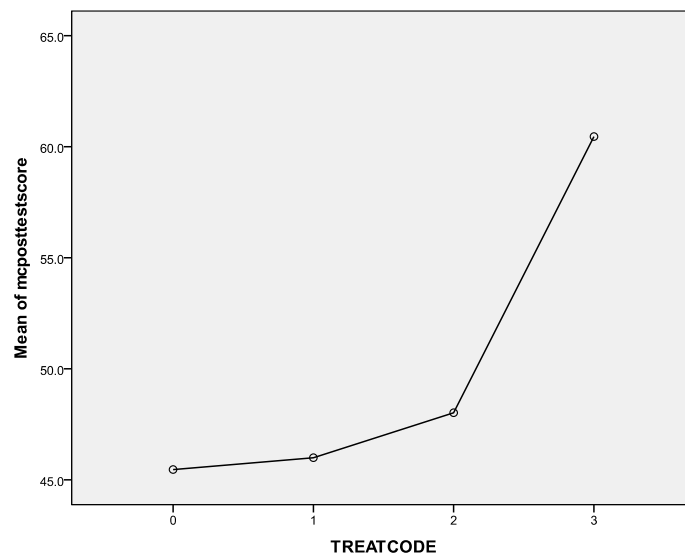**Figure 2. Multiple Choice Posttest Means by Treatment Group**

100

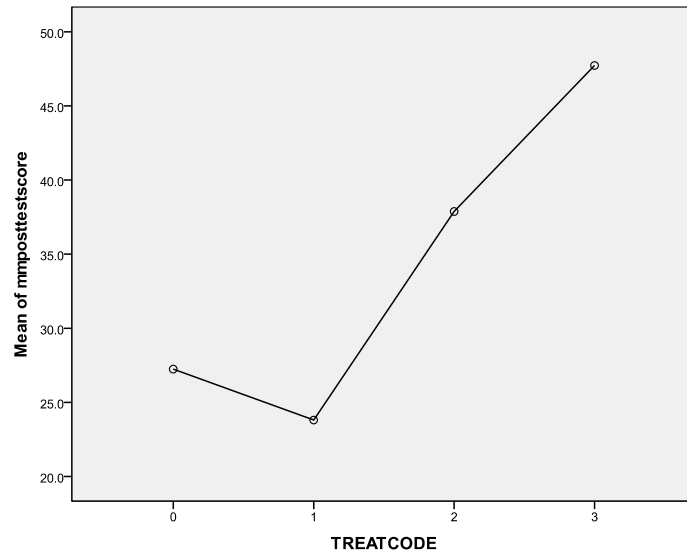**Figure 3. Mental Model Posttest Means by Treatment Group**

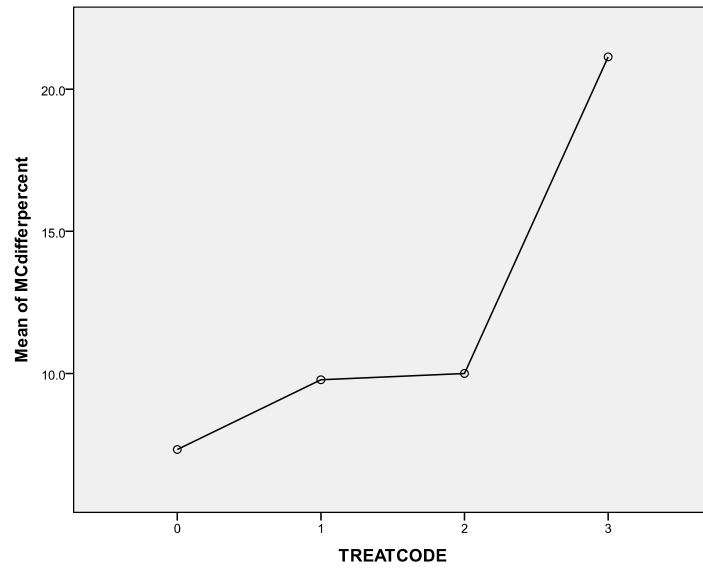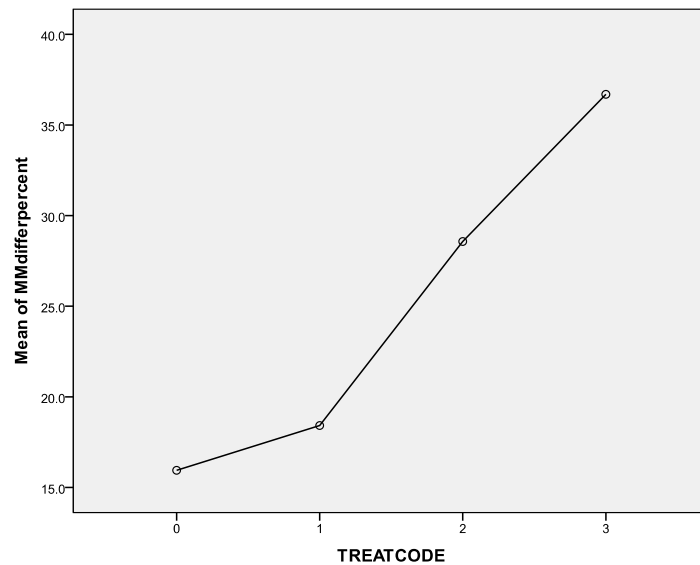**Figure 4. Multiple Choice Gains by Groups**

Figure 5. Mental Model Gains by Groups

Graphically, the difference between treatment group three's scores, the self-regulated learning group, stand out across both posttest and gain measures. Treatment group two's scores were also higher when compared with the control group and treatment group one, however as noted above, statistically significant differences were found only in the mental model posttest and gains. These results support the effectiveness of animation viewing employed while reading the text packet and to a lesser degree after reading the packet although not conclusively given this study's mixed findings in relation to group two. While treatment participants score more highly than the control group, treatment group one's scores are not significantly higher, and in the case of mental model posttest score, is actually lower than the control group's participants. This lower value was hypothesized to potentially be a function of participants who did not put forth a conscientious effort during the experiment. As can be noted in Tables 1 and 2, some participant scores actually decreased following the intervention. While these scores were

kept in the preceding analysis given the fact that animation viewing has been shown to result in no learning gains and in some instances increased confusion and lower posttest scores (Hegarty & Narayanan 1998; Libarkin et al. 2003), these negative change data were removed to evaluate their impact on treatment group differences.

In the first comparison, cases were selected for removal if a participants' multiple choice posttest score was lower than their pre-test score. This resulted in a sample size reduction from 175 participants to 155 participants. Model results, however, did not change. Treatment group three's scores were found to differ significantly from group one's scores on the multiple choice posttest, ($M = 13.7$, $SD = 4.3$), $F(3, 151) = 4.613$, $p = .004$. Using the multiple choice gain score as the dependant variable, treatment group three's scores were significantly higher than both group one and group two scores. The mean difference with group one was ($M = 8.1$, $SD = 2.2$), $F(3, 151) = 10.787$, $p = .000$ while the difference with group two was ($M = 7.4$, $SD = 2.3$), $F(3, 151) = 10.787$, $p = .000$.

The same selection criterion was applied to the mental model gains variable to determine if any group differences varied when removing those participants that did not comply with study expectations. In this case, the sample size was only reduced by seven participants from 175 to 168 and as the results indicate, no changes occurred. Treatment group three's posttest scores only differed significantly from group ones' scores, ($M = 27.3$, $SD = 5.2$), $F(3, 164) = 11.030$, $p = .000$ as was the case with the gains scores, ($M = 20.5$, $SD = 4.2$), $F(3, 164) = 12.373$, $p = .000$. Thus, after removing the participants whose posttest scores were lower than their pretest scores, no outcomes changed.

Next, posttest assessment scores and derived learning gains were examined to determine whether participants' content area prior knowledge, verbal ability and/or mental rotation ability impacted assessment performance. In this initial analysis stage, median splits were utilized to categorize participants into high and low prior knowledge, verbal ability and rotation ability groups for analysis. In later regression analyses, these variables are not dichotomized to utilize the entire scale of these pretest assessment measures as possible explanatory variables.

Prior knowledge median splits were constructed from both the multiple choice and mental model pretests and used to analyze both multiple choice and mental model posttest scores and derived learning gains for all participants. In each of the four posttest measures, high prior knowledge participants scored significantly higher than low prior knowledge participants. Using the multiple choice pretest split and the multiple choice posttest as the dependent variable, a one-way ANOVA confirms that high prior knowledge participants outperformed low prior knowledge participants $F(1, 173) = 145.63, p = .000$. This result was confirmed when replacing the multiple choice posttest with the mental model posttest, $F(1, 173) = 16.63, p = .000$. Similarly, using the mental model pretest median split and the multiple choice posttest score as the outcome, the same results were found, $F(1, 173) = 29.95, p = .000$ as well as when the mental model posttest score was selected as the outcome measure, $F(1, 173) = 55.84, p = .000$. As expected across prior science education research, participants with a greater level of prior knowledge obtained higher scores on both posttest assessments. To control for differences in participants' prior knowledge and their impact on posttest scores,

comparisons were examined using the calculated learning gains from the multiple choice and mental model assessments.

Learning gains were first compared by prior knowledge levels using the multiple choice pretest median split. No significant differences were found between the groups using both the multiple choice gain and the mental model gain variable. Similarly, no significant difference was found using the mental model pretest median split and the mental model gain score. However the high prior knowledge group as defined using the mental model pretest median did score significantly higher on the multiple choice gain variable, $F(1, 173) = 5.82$, $p = .017$. These results indicate that in nearly every case, prior knowledge did not play a significant role in the learning gains attributable to the experimental conditions.

Next, these analyses were repeated for the control group and the combined treatment groups to see whether prior knowledge levels differentially impacted posttest scores and learning gains. Results from the control group subset indicated that when high and low prior knowledge groups were defined by the multiple choice pretest, significant differences were found for the multiple choice posttest, $F(1, 41) = 39.662$, $p = .000$, and the mental model posttest, $F(1, 41) = 10.061$, $p = .003$, although no significant differences were noted in learning gains. Similarly, for the control group prior knowledge split using the mental model pretest, significant relationships were found for the multiple choice posttest, $F(1, 41) = 22.39$, $p = .000$, and the mental model posttest, $F(1, 41) = 46.48$, $p = .000$ but no significant differences were found between the learning gains variables.

Statistically significant differences between high and low prior knowledge participants in the treatment group, defined by the multiple choice pretest median split, were found between both posttest assessments, multiple choice $F(1, 131) = 115.882$, $p = .000$ and mental model $F(1, 131) = 9.97$, $p = .002$, however no differences were found between measures of learning gains. When the treatment group is examined using the mental model pretest median split, both posttest measures again have statistically significant differences, multiple choice posttest results are $F(1, 130) = 15.34$, $p = .000$ and mental model posttest results are $F(1, 130) = 29.163$, $p = .000$. However, unlike the preceding measure and the control group's findings, the multiple choice learning gain measure is found to differ significantly, $F(1, 130) = 5.17$, $p = .025$.

These results indicate that, in the majority of cases, participants' content area prior knowledge did not affect learning gains in any of the treatment conditions. The lack of difference may be explained by the fact that most participants in this study exhibited low levels of prior exposure to the content area, and thus very low pretest scores. Therefore the median splits, which occurred at zero percent for the mental model pretest and 35% for the multiple choice pretest effectively grouped participants with minimal prior understanding levels with those participants with the high levels of prior knowledge groups. Figures 6 and 7 present the distribution of prior knowledge scores. As is abundantly clear on the mental model pretest figure, without the prompts and educated guesses available on the multiple choice exam, most students had little content information available in prior memory to complete open-ended and diagrammatic questions. One potential solution was to alter the level at which high prior knowledge was defined, for example using 75th percentile instead of the 50th percentile. However,

because this dichotomization technique is not common in prior animation research and because individual score differences can be better modeled in multiple regression analysis, this analytic technique was not pursued.

Figure 6. Pretest Multiple Choice Scores by Group
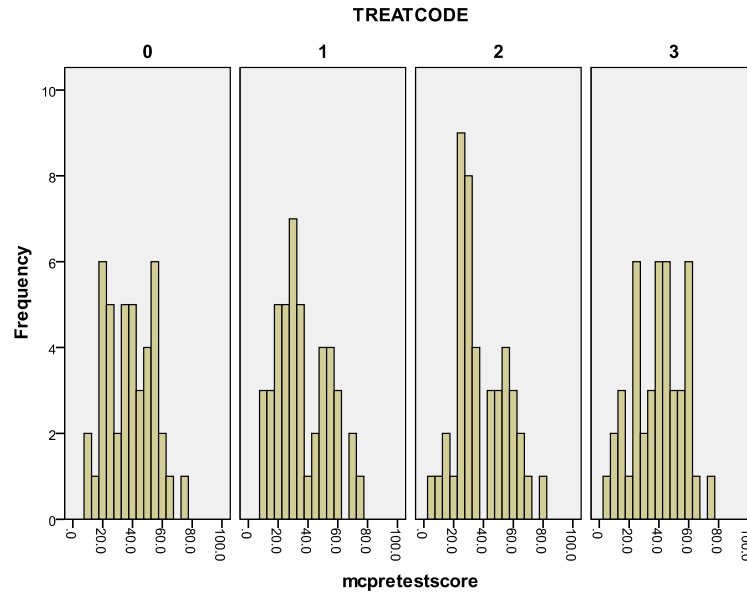
Figure 7. Pretest Mental Model Scores by Group
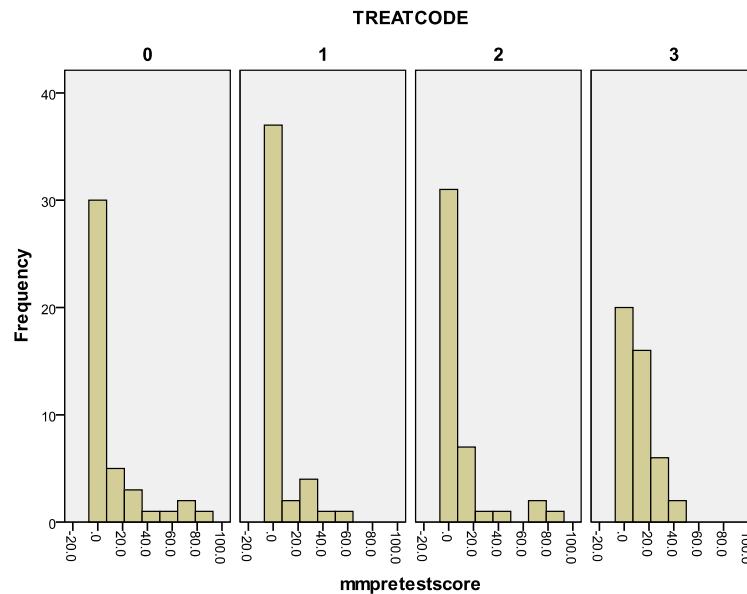
Next, verbal ability groups were compared across multiple choice and mental model posttest scores and the calculated learning gains. High verbal ability participants, on average, had better scores on both the multiple choice and mental model posttests (Figures 8 and 9), however only the multiple choice posttest difference was statistically significant, $F(1, 173) = 4.82$, $p = .030$. Verbal ability grouping was not found to be statistically significant for either learning gain measures although higher scores were observed by high verbal ability participants (Figures 10 and 11).

Figure 8. Verbal Ability and Multiple Choice Posttest Scores
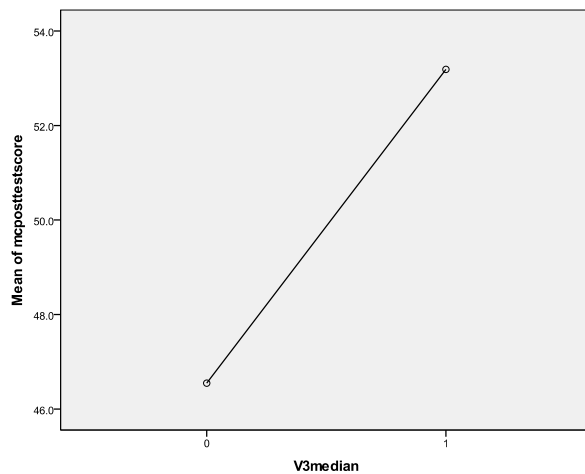
Figure 9. Verbal Ability and Mental Model Posttest Scores
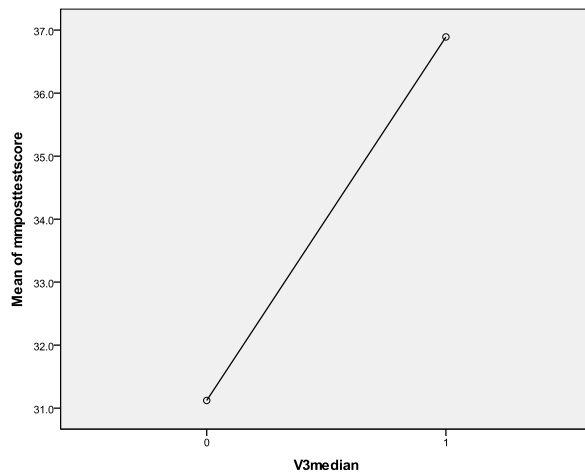
Figure 10. Verbal Ability and Multiple Choice Posttest Gain Scores
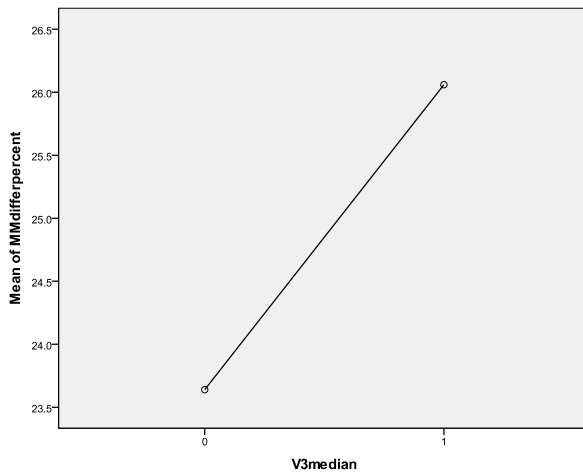
Figure 11. Verbal Ability and Mental Model Posttest Gain Scores
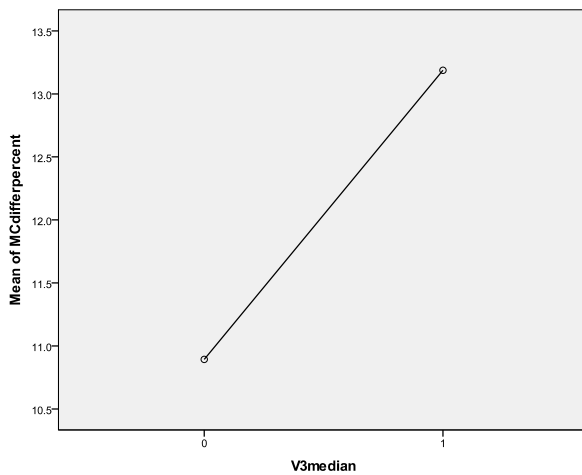
Verbal ability groupings were further analyzed to compare for differences within the control and treatment groups. No significant differences in posttest scores and learning gains by verbal ability were found for the control group while a single significant difference was observed for high verbal ability participants in the treatment group on the multiple choice posttest, $F(1, 131) = 5.575$, $p = .020$.

While text comprehension, as measured by the verbal ability test, improved posttest performance, comparative differences between treatment conditions, for the most

part, were not statistically significant. That these differences were only observed in the multiple test measures rather than the mental model measures implies that verbal comprehension aided the selection process of answers when provided however did not prove beneficial given the diagrammatic understandings and representations required for in the mental model assessment.

Next, posttest scores and learning gains were compared by mental rotation ability groups. A significant difference was found for high rotation ability participants and the mental model posttest score, $F(1, 173) = 10.44$, $p = .001$ while no significant difference was discovered for the multiple choice posttest. This pattern repeated itself for the multiple choice and mental model gain scores. Mental model gains were greater for high rotation ability students and significant at the 0.05 level, $F(1, 173) = 3.98$, $p = .048$, however the gain was not significant for the multiple choice gain. Lastly, within group rotation ability differences were examined for the control and treatment groups. No significant differences were discovered for the control group across posttest and gains measures while a single significant difference was found for the treatment group. The difference existed for the mental model posttest score where high rotation ability participants in the treatment group scored significantly higher than the low ability participants, $F(1, 130) = 7.196$, $p = .008$.

These results, while generally inconsistent across the posttest and gain measures, seem to indicate that those participants with a greater ability to visualize figure transformations were more likely to encode relevant content from the animations and generate appropriate diagrams for assessment by the mental model posttest. However, this improved conceptual understanding did not translate into declarative knowledge

gains as evaluated by the multiple choice tool. The preceding analyses sought to answer the study's primary research questions and illuminate significant relationships between the measured explanatory variables and the dependent assessment variables. While preliminary differences between treatment and control group performance and participant characteristics were identified, these univariate analyses did not seek to determine which variables exhibited the greatest influence in explaining learning differences across intervention conditions. Moreover, the preceding analysis did not look at possible interactions among the explanatory variables seeking to explain and predict the dependent variables. These considerations are addressed in following regression analyses section.

Section 5. Regression Analysis

Regression equations were constructed to determine which independent variables and combinations of independent variables significantly explained measured variability in posttest scores and learning gain outcomes. Because the outcomes were measured on two scales, continuous level data on the multiple choice posttest and both gain scores and ordinal data for the mental model posttest, two different types of regression modeling were employed. For the multiple choice posttest measures and the learning gains scores, linear regression models were selected while odds probability (OP) ordinal regression was used for the ordinal mental model levels.

The first linear regression models examined participant posttest and gains variability explained by the student ability variables which included the V-3 vocabulary test and Vandenberg and Kuse's Mental Rotation Test as explanatory variables. Vocabulary scores were found to significantly predict multiple choice posttest scores, $\beta =$ .35, $t(173) = 2.48$, $p = .014$ however only a small proportion of variance was explained, $R^2 = .034$, $F(1, 173) = 6.15$, $p = .014$. When the multiple choice learning gain variable was selected as the outcome measure, neither the vocabulary test nor the mental rotations variable were found to significantly explain multiple choice learning gains.

After determining that the vocabulary test was a significant, albeit, weak explanatory variable, a second model was constructed incorporating the multiple choice pretest score as a second block in a hierarchical stepwise regression model to determine the variability explained by participant prior knowledge. The resultant model found that the multiple choice pretest score was significant, however, at the expense of the vocabulary score. Thus the variability explained by the vocabulary test measure was found to be better explained by the prior knowledge pretest. Moreover, the prior knowledge score explained 63.5% of post score variability, $\beta = .95$, $t(173) = 16.83$, $p = .000$ and $R^2 = .635$, $F(1, 173) = 149.65$, $p = .000$. The vocabulary score model parameters changed to $\beta = .09$, $t(173) = 1.05$, $p = .295$ in the model.

Next, the treatment conditions were entered into a hierarchical regression model, with the prior knowledge pretest score as the first block, to determine whether an intervention type explained variability in the posttest measure over and above the prior knowledge score. Only treatment group three, the self-regulated learning intervention,

was found to be significant and the variable was found to explain an addition 6.9% of the

variability in the post test score (Table 5).

Table 5. Summary of the Hierarchical Regression Analysis for the Prediction of the Multiple Choice Posttest

Predictor Variable	R^2	ΔR^2	β	Standard Coefficient (beta)	t	Sig.
Step 1 MC pretest	.633	.633	.955	.795	17.26	.000
Step 2 Treatment 3	.702	.069	12.20	.263	6.30	.000

Note: Number of cases 175.

The final model using the multiple choice posttest score as the dependent variable

incorporated the prior knowledge variable and treatment group three variable as separate

blocks in a hierarchical regression and added a third block including the student

characteristics reported on the demographic instrument. These block three variables were

entered stepwise into the model and one of the measures, self-reported GPA in the major,

was significant over and above the prior knowledge and self-regulated learning treatment

(Table 6). The inclusion of the GPA major added 3.3% to explanatory power to the

model.

Table 6. Summary of the Hierarchical Regression Analysis for the Prediction of the Multiple Choice Posttest

Predictor Variable	R^2	ΔR^2	β	Standard Coefficient (beta)	t	Sig.
Step 1 MC pretest	.404	.404	.723	.692	10.021	.000
Step 2 Treatment 3	.492	.088	10.311	.322	4.637	.000
Step 3 GPA major	.525	.033	5.016	.184	2.684	.000

Note: Number of cases 175.

The same approach was considered for the multiple choice learning gain score

and the mental model learning gain score, since both of these variables were measured on

a continuous scale. As noted in the prior section, neither the vocabulary nor mental

rotation variables were found to significantly predict posttest gain scores. When the

treatment variables were subsequently entered into the hierarchical model as the second

block, only treatment condition three, the SRL animation group, was significant, $\beta=$

12.09, $t(173) = 6.24$, $p = .000$ and $R^2 = .184$, $F(1, 173) = 38.93$, $p = .000$.

The last model constructed included an additional block containing student

demographic variables. As with the preceding analysis examining only the posttest score,

the gain score's variability was further explained by the inclusion of major GPA variable.

Model results are presented in the table 7 below.

Table 7. Summary of the Hierarchical Regression Analysis for the Prediction of the Multiple Choice Gain

Predictor Variable	R^2	ΔR^2	β	Standard Coefficient (beta)	t	Sig.
Step 1 SRL Treatment 3	.174	.174	11.85	.443	5.09	.000
Step 2 GPA major	.223	.049	5.09	.223	2.56	.012

Note: Number of cases 175.

Mental model gain scores were then analyzed using the same stepwise,

hierarchical regression procedure. Unlike the multiple choice gains model, the

Vandenberg and Kuse Mental Rotation score was a significant predictor, $\beta= .18$, $t(173)$

$= 2.57$, $p = .011$ and $R^2 = .037$, $F(1, 173) = 6.60$, $p = .011$ in the single block model.

When the treatment conditions were added in block two, both treatment three, the SRL

114

group, and treatment two, the animation after group, were significant at the expense of the MR rotation score (Table 8). Thus variability initially explained by rotation score differences was better explained by the two significant treatment conditions. No demographic variables were significant when the third block was added to the model.

Table 8. Summary of the Hierarchical Regression Analysis for the Prediction of the Mental Model Gain

Predictor Variable	R^2	ΔR^2	β	Standard Coefficient (beta)	t	Sig.
Step 1 MR score	.037	.037	.13	.14	1.96	.052
Step 2 SRL Treatment Animation After Treatment	.154	.117	18.43	.36	4.73	.000

Note: Number of cases 175

Section 7. Ordinal Regression

The ordinal regression analysis used the proportional odds (PO) model for the dependent variable mental model posttest score. Figure 13 below graphically displays the posttest mental model differences by treatment condition. Visually, it is apparent that participants in treatment group two and three produced higher scores on the mental model posttest tool given their overall low number (i.e., cumulative frequency) of scores on the ordinal scale's low end (i.e., mental model levels 0, 1 and 2).

Figure 12. Cumulative Frequency of Posttest Scores by Treatment Condition
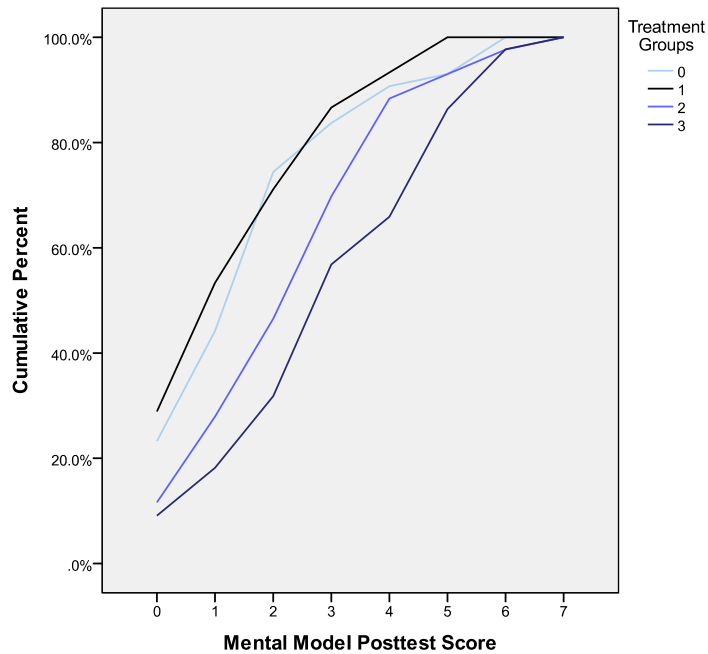
In the ordinal regression analysis, the logit link function was used for the model because it is appropriate for analyzing ordered categorical data when the observations are relatively evenly distinguished across all categories. In the model, the odds of the event occurring are defined as the ratio of the probability of event occurrence to the probability of the event not occurring. Moreover, it's a cumulative logit model, or proportional odds model, because the cumulative response probability calculations compare the prediction of inclusion in one category or lower given the known value of the explanatory variable (Walters et al. 2001). For example, the first coefficient in the model would compare the probability of the observation being in category one versus the cumulative probability of falling into all remaining categories.

The first model included the V-3 vocabulary test and the Vandenberg and Kuse Mental Rotation test as predictors. Results from the model-fitting test indicated that the

116

model with predictors outperformed the model without these pretests as predictors χ^2 (1,

$N = 175) = 81.41, p = .001$.

However, the null hypothesis was not rejected for the test of parallelism, χ^2 (1, N

$= 175) = 544.11, p = .000$, indicating that the coefficient varied across each mental model

level, a violation of a key assumption in ordinal regression, and thus a multinomial

regression should be considered.

Multinomial regression results, however, did not find that the inclusion of the V-3

vocabulary test and Vandenberg and Kuse's Mental Rotation produced a better model

than a model with no predictors, χ^2 (1, $N = 175) = 536.01, p = 1.00$.

After removing these pretest measures given their inconsistent ability to predict

mental model levels, the second model included all treatment conditions aggregated to

one group for comparison with the control group. Results from this model indicated that

this predictor, inclusion in a treatment condition, outperformed the model with no

predictors, χ^2 (1, $N = 175) = 4.38, p = .036$.

Moreover, the test of parallelism assumption was met; therefore coefficients could

be examined to determine the impact of treatment on mental model levels. The

coefficient was negative, -.644 for the control group, indicating that participants in this

group was 1.9 times less likely to produce a high score on the posttest.

Given that participation in a treatment condition did result in an improvement of

model fit, dummy coded variables for the individual treatment conditions were entered

into the model to determine which treatment condition or conditions increased the

likelihood for higher levels of mental model posttest performance. The resultant model

was a better predictor than the intercept only model, χ^2 (1, $N = 175) = 24.46, p = .000$ and

both treatment group two, animations viewed after reading, and group three, animations

viewed during reading, were found to increase the probability of higher performance on

the mental model posttest. Treatment group two's coefficient,.828, indicated a 2.3 times

odds of producing a higher mental model score while treatment group three's coefficient,

1.506, indicated that participants in this group had a 4.5 times likelihood of producing a

higher posttest score. While ordinal regression does not provide a true R^2 explaining the

percentage of variation observed in the dependent variable, pseudo-R^2 values have been

devised and presented below for the model. If the chosen predictors for the model are

effective, pseudo R^2 scores greater than zero will be calculated.

Table 9. Pseudo R-Square model output for treatment groups

Cox and Snell	.130
Nagelkerke	.133
McFadden	.037

Link function: Logit.

A final model was constructed supplementing the two significant treatment

conditions with coded demographic characteristics to determine whether model fit could

be improved. The resultant model was significant, χ^2 (1, $N = 175$) = 89.147, $p = .000$ and

increased the pseudo R^2 with the inclusion of several additional predictor variables.

Table 10. Pseudo R-Square model output for treatment groups and demographics

Pseudo R-Square	
Cox and Snell	.565
Nagelkerke	.577
McFadden	.215

Link function: Logit.

Variables found to improve model fit included gender (-1.562, p = .002), academic years one (-3.375, p = .034) and two (-3.025, p =.003), all levels of overall GPA (18.584, p = 000), math courses greater than two (2.040, .008), and course attitude greater than 4 (-1.408, p = .027). In interpreting the results, the sign of the estimate is important given that negative values indicate the likelihood of producing lower mental model levels. Because male participants were coded one and female participants zero, results indicate that female students were more likely to produce higher level mental models if they were participants in treatment groups two or three. Freshmen and sophomore were found to produce lower level mental models in comparison with juniors and seniors as were students indicating higher positive attitudes toward the course. The likelihood of significantly higher mental model levels was also identified in participants with high overall GPAs and more mathematics coursework.

Section 8. Qualitative Procedural Data Analysis

While the preceding analysis inferred cognitive processes to explain differences between treatment conditions and resultant performance on the posttest assessments, procedural or process data were collected from twelve randomly selected participants in an effort to identify specific learning strategies employed and whether differences existed between treatment conditions. Three participants from each treatment condition including the control group were digitally recorded while reading the textbook packet and while examining the animations with the exception of the control group members. These participants were instructed to vocalize their thoughts during the learning episode and were prompted with the question, "What are you thinking?" if they remained silent for

more than 45 seconds. One-hundred and ninety-two minutes of audio data were collected and transcribed by the study's author producing twenty one pages of text or 6034 total words ($M = 593$ words per participant). The initial transcription and audiotapes were compared by a geosciences faculty member to verify the accuracy of the transcription and to edit any missing utterances and/or erroneous interpretations in the text document. These data were then coded according to Azevedo et al.'s model (2004) of Self Regulated Learning which was based on Winne (2001) and Pintrich's (2000) segmentation of regulatory processes into four behavioral phases: 1) planning which includes the activation of prior knowledge and setting goals through the coordination of operations during the learning exercise, 2) monitoring which are metacognitive assessments of content understanding , 3) strategy use which are the specific learning strategies employed in order to comprehend the knowledge representations provided in the intervention, and 4) task difficulty and demand which entails the learner evaluating the content in relation to their current understanding and the learning goals to intentionally control behavior. Azevedo et al. (2004) formulated and described sub-processes within these four areas specific to learning in hypermedia which were adapted to the animation setting used in this study. Appendix C provides a reproduction of the processes, descriptions and examples developed by Azevedo et al. (2004).

Transcribed think aloud data for each participant was segmented in the text file to align to variables described by Azevedo et al. (2004) and coded using this framework for analysis and group comparisons. Coded results are presented by frequency and treatment condition in Table 9 below and only include variables observed in participants' transcripts.

Table 11. Number and proportion of self-regulated learning variables employed by treatment group

Class & Variable	Control (N = 3)	Animation Before (N = 3)	Animation After (N = 3)	SRL Animation (N = 3)	Raw frequencies
Planning					
Goals	0	0	3	0	3
Prior knowledge activation	6	5	8	5	24
Monitoring					
Judgment of learning	9	11	16	6	42
Feeling of knowing	12	10	19	10	51
Self-questioning	0	0	0	1	1
Content evaluation	0	1	0	1	2
Identify adequacy of information	0	0	0	1	1
Strategy use					
Selecting a new informational source	1	2	2	1	6
Coordinating informational sources	1	1	5	3	10
Read a new paragraph	1	1	2	0	4
Summarization	17	21	26	6	70
Rereading	1	0	3	0	4
Inferences	0	0	1	0	1
Hypothesizing	0	0	1	0	1
Knowledge elaboration	1	1	2	2	6
Analogy use - new	0	1	1	0	2
Task difficulty and demands					
Help-seeking behavior	4	2	6	8	20
Task difficulty	1	0	1	2	4
Control of context					
Interest					
Interest statement	1	6	1	1	9

As is evident across all groups and in total, the most frequently employed behavior observed during the learning episode was the summarization strategy. Because most participants lacked sufficient background knowledge about atmospheric pressure in general and Hadley Cell circulations in particular, the most frequent vocalizations were either direct quotations or paraphrases of content presented in the text. Often, these

utterances were accompanied by a follow-up monitoring process, for example a judgment

of learning or a feeling of knowing as exemplified in the exemplar segments presented

below.

> ”A knot is the speed of one nautical mile per hour – I never knew that. 1.5 statute
> miles per hour…so…hmm…gotcha…” SRL group – participant #2

> "So the Trade Winds feed into it and the Anti-Trade Winds go away from it
> Umm…so the STHs are the anti-cyclones...umm I get it" Animation After –
> participant #1

These follow-up monitoring processes were the second and third most frequently

observed behaviors among the subjects and in most cases, the segments took the form of,

"I didn't know that" or "I've never thought about it" judgments of learning or an "OK

that makes sense" a feeling of knowing. As the counts in Table 9 indicate, participants

expressed more statements of understanding in relation to the learning materials than

confusion (i.e., 51 segments vs. 42 segments) although considerable confusion was

expressed by several participants when examining the isobar diagrams included in the

learning packet. A reproduction of one of the diagram follows as do exemplar statements

made when examining these diagrams.

Figure 13. Global Average Pressure Differences by Seasons
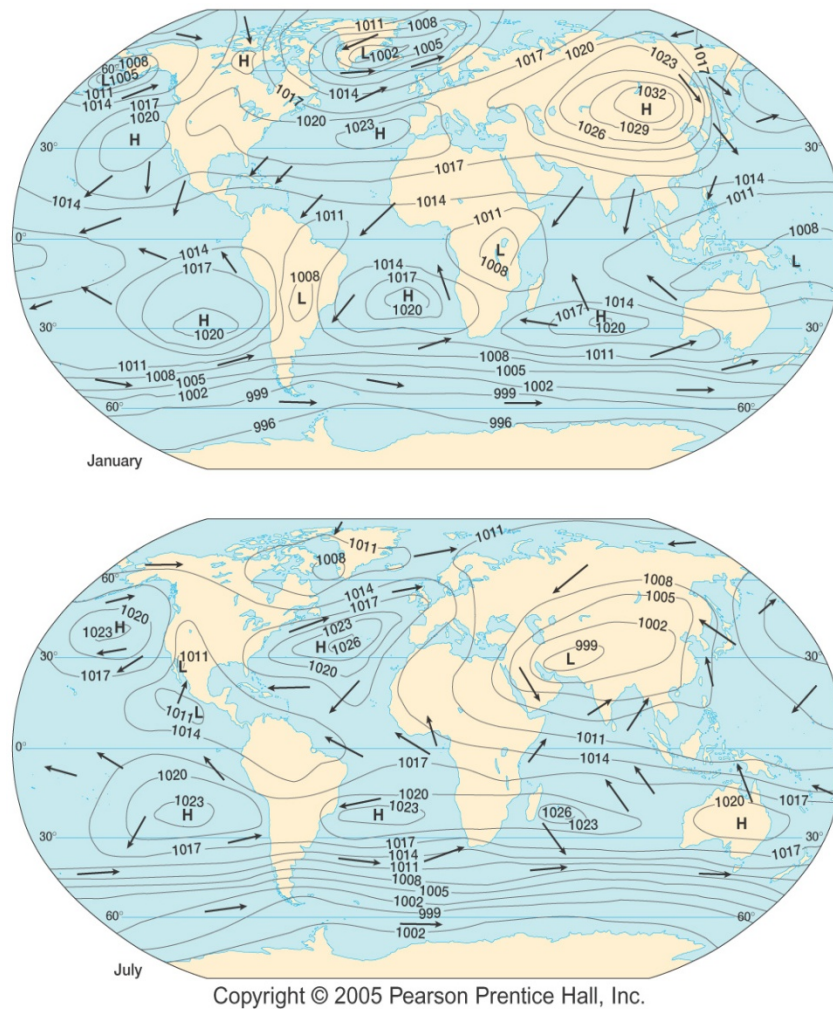
"I have no idea what this means." (Animation After - participant #2)
"The figure…looks like it's really difficult to read." (SRL participant - #1)
"Honestly I wouldn't probably look at this if I was reading the book because I don't even know what I'm looking at." (Animation After - participant #1)

Prior knowledge activation was the next most frequently employed behavior however, based on the participant's segments, most had simply heard the term or terms before and did not possess any real depth of knowledge about the concept. For example, SRL participant #3 stated, "I remember the pressure gradients" but did not offer any additional description or explanation implying simple term recognition. The same

interpretation can be assumed from Animation After participant 1's statement, "Umm equatorial front doldrums…weak horizontal air flow, erratic winds, low pressure instability in Hadley cell…yes I knew a little about that" and animation after participant 2's comment, "I haven't heard this since my high school class." Both statements reveal a less than convincing understanding of these target concepts.

Two subjects, both science majors, seemed to possess more robust prior knowledge concerning aspects of the target content based on the following statements, however qualified their depth of understanding. For example, "I've heard of all this stuff before….I just didn't connect it" (Animation After - Participant #2) and "I know a little bit about the Coriolis effect from my physics class…I don't really remember much of it though" (Animation Before - participant #3). Both of these qualifying statements indicate metacognitive monitoring of their understanding and recognition that their prior understanding of the topic or exiting mental model was incomplete.

The final variable observed with relative frequency (i.e., 20 instances) was help-seeking behavior. In all cases, this code indicated that participants asked a direct question of the researcher for content clarification given their inability to comprehend a concept or process described in the text packet, by diagram or in an animation. This behavior was indicative of an individual's assessment of task difficulty and was noted by subjects across each treatment and the control group. Because the researcher was also their course instructor, many of these questions occurred simply due to their familiarity with typical laboratory procedures where discourse is encouraged and questions are answered. However, during the experiment, these questions were uniformly greeted with the response, "I'll be happy to answer any question and/or clarify any topic after the

completion of the experiment. Please make note of any question on the text packet" so

that no participant received unequal aid thereby impacting posttest assessment

performance. Examples of typical questions were:

> "Which way is convergence?" (SRL group - participant #3)
> "What are isobars?" (Control group - participant #2)
> "Is the blue low pressure and the red is high pressure?" (Animation After group - participant #2)
> "What does it mean as shown in picture?" (SRL group - participant #1)

As noted by the similarity of enacted behaviors across the groups, little variation

was discovered with the sole exception of a single specific strategy use, the coordination

of informational sources. This variable was identified and described by Azevedo as an

important cognitive process that occurs when learning with multiple representations in

hypermedia environments. And while this study's setting was not a true hypermedia

environment, the availability of the animations on the laptops offered learners the

opportunity to create linkages between the available representational forms. Both the self

regulated learning group and the animation after group, participants found to produce

higher posttest scores and learning gains, verbalized this strategy use with more

frequency than the control group and the animation after group. Moreover, as can be

noted from their statements, the representational coordination was often relating text

passages or static diagrams to processes observed in animations. For example:

> "When reading what an anticyclone is and then looking at the screen it helps me understand it a lot more than what they are trying to explain to me." (SRL group - participant #1).

> "The figure on 5.15 looks like it's really difficult to read but on the computer seeing all the movements it's a lot easier than just all the lines on the figure." (SRL group - participant #1)

These comments also indicate that the process of coordinating these representations can initiate metacognition whereby the participants evaluate their understanding or even their misconceptions based on a single representation source. Whereas target content confusion may continue to exist in those groups where participants are not able to coordinate the text and diagrams concurrent to animation examination, these participants recognized their own surficial or erroneous understanding and sought confirmation through comparison with an available, alternate representation.

When examining SRL behaviors specific to animation viewing, several problems which affected information extraction from the animations were noted as were several affordances. The problems identified primary encompassed misunderstanding symbology and/or not being able to comprehend the differences in circulation being represented in the animations. For example, SRL group - participant #2 asked," What really is the difference between the cyclone and the anti-cyclone because they are both moving in the same direction aren't they?" Another example of confusion was stated by Animation After - participant #2, "So there are bunch of arrows spinning in different directions…hmm. " These spinning arrows and the colors assigned to the spinning arrows in the Hadley circulation also created confusion in Animation After - participant #3 who said, "I just can't tell if the arrows are pressure or temperature."

However, this confusion was often clarified after repeatedly viewing the animations. SRL group - participant #2, after not being able to differentiate cyclones and anti-cyclones initially, began to see the difference and followed up with the statement, "So anti-cyclones move clockwise and cyclones move counterclockwise." Animation After - participant #3 admitted confusion between cyclones and anticyclones after

reading the text and examining the static diagrams and used the animations to clarify understanding. This progression can be seen in the following verbalization sequence, "(I'll) try cyclone first because I'm not sure I understand the difference yet…So the cyclone starts at the ground and moves up… the anticyclone starts in the sky and comes down." This sequence clearly portrays the beneficial role animations can play with a participant who is serious about understanding the target content. This sentiment was repeated by SRL group - participant #1 who stated, "It helps to see the figures in motion on screen versus the paper." The ability to examine the animation sequences also elicited an understanding of atmospheric circulation difficult to convey by text and static diagrams alone. Animation After - participant #2 stated, "Interesting to see that a lot of movement of air is up and down as well." This observation is critical to understanding the cause and patterns of precipitation, and the lack thereof, explained by Hadley cell circulation and are often missing from novice learners' mental models. These vertical motions were also explicitly represented in the Hadley Cell animations and linked to the high and low pressure systems situated in the tropical latitudes. The following participant sequences exemplify student processing of these complex and dynamic relationships with varying degrees of success. "And then as it goes outward it cools off and then it comes inward it gets hotter I guess you can say… on either side of the equator. And then it gets to the equator and flows outward and gets cool… comes back in gets hotter in a continuous process." (Animation After - participant # 2)

> "So this is showing Hadley Cell circulation…and it appears that at the equator...warm air is convected (sic) towards to the equator and rises into the air and becomes cooling air and um I guess travels in the upper atmosphere in an arc sort of a pattern and then falls back down to the earth around the Tropics of

Cancer and Capricorn in NH an SH hemispheres respectively…pretty cool…I had no idea this was going on" (Animation Before - participant #3)

"Wind is pushed like into the equator and then out I guess to I guess 30 degrees latitude or longitude...It's got something to do with the low and high pressure…the low is in the middle and the high." (Animation Before - participant #2)

These sequences exemplify the animations' ability to link temperature, pressure and circulation; however, in each case important details and relationships are missed in these vocalizations. For example, while the first and second sequences explicitly link these movements to temperature, the third segment ignores temperate depicted by colored arrows in favor of the high and low pressure symbols obviously ignoring their interrelationship. Thus in each case, an important aspect of the circulation system was omitted and therefore resulted in incomplete mental model representations.

One participant recognized the difficulty of comprehending the entire system using only the three representation forms offered and believed additional understanding was possible with a scaffold typically available during formal learning in the environment where the study was conducted. SRL group - participant #1 stated, "what would really help me more is like if a teacher were teaching with the moving diagram instead of just reading and looking at the diagram (sic: animations)." Thus while the animations offered affordances to some learners, others realized the limitations of the representational form and desired an additional learning aid that would facilitate their content understanding. And while the self regulated learning group and animation after group tended to realize greater gains in system understanding as measured by the quantitatively analyzed posttests, few SRL process differences were noted in the process data. Thus these gains are most likely explained by text and diagram examination generating requisite prior

knowledge that enabled these participants to selectively extract salient information from the animations.

Chapter 5: Discussion

The preceding analyses quantitatively and qualitatively evaluated this study's two main research questions. Results indicated that the inclusion of animations in a treatment condition did not necessarily lead to increased content area understanding in comparison with the control group. Statistically significant differences were not discovered between each treatment group's posttest scores, as measured by either declarative knowledge in multiple choice posttest or deeper understanding in the form of the mental model posttest. Thus the ability to view animations in and of itself was not found to increase Hadley system understanding. These findings corroborate prior research finding no significant advantage for the use of animations for learning (Hegarty & Narayanan 1998; Lowe 2005). However, when evaluating the differences between treatment groups and the control group individually, significant differences were found between the Self Regulated Learning condition (i.e., treatment group three) on both posttest measures and the animation after condition (i.e., treatment group two) on the mental model posttest. These findings suggest a sequence effect for the timing of animation viewing to increase learner understanding which is contrary to a similar study which included animations and video presentations (Velazques-Marcano et al 2004). Thus, when animations are viewed by the learner at a specific time in the learning activity, increased understanding of Hadley Cell circulations resulted.

The identification of an optimal time period in which to examine the animations can be understood through prior research identifying the affordances offered by dynamic, external representations. First, these representations take advantage of the multiple channels of sensory input and thus provide more content information in a given time

period for processing within working memory (Baddeley 1992; Clark & Pavio 1991). Moreover, viewing the additional representations allows the learner to evaluate current understanding against the depicted processes enabling the use of the metacognitive strategy of representational coordination (Green & Azevedo 2007; Mayer 2001). However, requisite minimal levels of prior knowledge must be met by the learner before these affordances can affect system understanding. In this study, this condition appeared to be met by the animation after reading and the animation during reading treatment groups. Prior knowledge of the system depicted in the animations has been shown to allow for top down processing rather bottom up processing. When learners engage in top down processing, they are able to attend to the salient features and interrelationships occurring within the dynamic visualization (Goldman 2003; Lowe 2005; Ploetzner et al. 2005). Novice learners without an appropriate level of prior knowledge engage in bottom up processing and are frequently drawn to seductive details or visually enticing actions or events that may not portray the most important processes in the depiction for understanding (Kriz & Hegarty 2007). The necessity of system prior knowledge is especially important before viewing animations of a complex system like Hadley Cell circulation for this study's sample (Schnotz and Rasch 2008).

As these results indicate, the paucity of prior knowledge related to the content area reduced the effectiveness of viewing the dynamic representations, since most participants had no systematic guidance to focus their attention to encode relevant information. Thus, when participants examined the animations prior to reading the textbook packet containing verbal explanations and diagrams, these learners were unable to evaluate the relative importance of elements being displayed in the cyclonic, anti-

cyclonic and Hadley Cell circulation patterns. For example, learners were not able to decipher representational symbology, like the meanings implied by arrow color and/or directionality in the horizontal or vertical plane, and the important interrelationships displayed between temperature, pressure and wind in the cumulative Hadley Cell representation. However, when the animations were viewed while examining the text and diagram descriptions (i.e., the SRL condition) or after examining these representations in the packet, most participants were better able to 1) understand the symbology depicted in the representations (e.g., red arrows indicate high temperatures and blue arrows indicate lower temperatures and the letter H indicated high pressure and the letter L indicated low pressure) 2) understand the causal mechanisms responsible for the different temperature and pressure conditions and their geographic positions and 3) understand the interrelationships between the temperature and pressure systems and the resultant wind and weather conditions depicted in the Hadley Cell animation. While these circulation systems remained relatively complex to these novice learners as evidenced by the range of posttest scores reported in Tables 1 and 2, participants in these two treatment conditions groups were found to produce greater posttest scores inferring that they were better able to identify key frames and encode relevant information presented in the animations.

While encoding differences based on prior knowledge appear to have affected system understanding, the animations selected for this study included design components that have been found to offer affordances for learners once a sufficient level of understanding was attained. The importance of design interactivity and the capability of the learner to pause, rewind and continuously loop the animations at user determined key

times in the sequences provided for the ability to match working memory limitations with

presentation information processing. This affordance has been noted in prior research

(Bodamer 2004; Lowe 2005) and can be understood by the piecemeal manner in which

learners have been shown to construct mental models from animations in eye tracking

studies (Hegarty 1998; Lowe 2008). In treatment group three, learners were able to

coordinate their reading and the diagrams with the animations through design

interactivity therefore exemplifying the temporal contiguity principle described in

Mayer's Theory of Multimedia Learning (Mayer 2001). Researchers have found that

animations often run at fixed rates and portray all system functions and interactions

simultaneously, overwhelming the novice learner's working memory sensory inputs

because of the high level of intrinsic load given system complexity and the lack of prior

knowledge (Chandler & Sweller 1991; Lowe 2008). While the animations selected for

this study could not affect intrinsic cognitive load, extraneous load was considered and

minimized through the use of bare bones animations containing few "bells and whistles"

or non-salient, seductive elements. The selection criteria can also be understood to align

to the coherence principle for design described for multimedia understanding (Mayer

2001). By selecting parsimonious representations with little extraneous information or

extensive detail overwhelming the level of understanding desired in the learner, design

decisions thus support noted cognitive processing bottlenecks facilitating learner's

understanding.

Additional statistical relationships and explanatory models were developed and

examined beyond the study's primary research questions, however few supplementary

independent variables were found to consistently increase the explanatory power of the

regression equations over and above the treatment condition. While high prior knowledge as measured by the pretest assessments resulted in better posttest scores, when the change in learning or learning gains were selected as the dependent variable, prior knowledge scores were not found to be significant across each comparison in the treatment groups. The exception to this finding occurred when prior knowledge was defined by the pretest mental model level and learning gains were measured by the multiple choice posttest. However, as described in the prior section, the median split for high prior knowledge by mental model placed any participant scoring greater than zero in this group. Thus any knowledge of the system, no matter how shallow and undifferentiated by knowledge type resulted in increases in declarative knowledge as measured by the multiple choice posttest.

Although inconsistent, this finding further supports the importance of top-down processing to take advantage of the affordances offered by viewing the animations (Kriz & Hegarty 2007). Participants with some prior exposure to the atmospheric system may have been able to metacognitively evaluate their limited understanding through the prompt to externally represent their mental model on the comprehensive assessment. Therefore, these individuals were able to employ strategies to achieve these goals through the examination of the textbook packet and produce learning gains during the treatment. In contrast, students with absolutely no prior knowledge of the system were less successful in monitoring their learning thus resulting in lower posttest scores. However, since high prior knowledge did not explain learning gains and posttest performance over and above the significant treatment conditions, the self regulated learning group and to a

lesser extent the animation after learning group, these conditions can be assumed to have benefited all participants rather than just one group based on the median split.

Similarly, verbal ability and mental rotation ability were not found to offer any more explanatory power to models including the treatment condition. While high verbal ability participants produced higher posttest scores for declarative knowledge (i.e., the multiple choice assessment tool), this pattern was not repeated for learning gains. Therefore these results may be interpreted to mean that while high verbal ability enabled these participants to outperform their peers on the pretest, the treatment conditions did not offer these participants any significant advantage in understanding over their low verbal ability peers and therefore did not result to greater gain scores.

A similar pattern was discovered for participants exhibiting a greater ability to perform mental rotations. While these participants produced higher posttest scores on the mental model assessment, no differences were found for learning gains. Mental rotation ability has been found to improve mental animations of mechanical systems from dynamic representations (Hegarty & Kozhevnikov 1999) and may explain why these participants were able to produce better comprehensive representations in their mental model posttests. However since no significant differences in learning gains were noted, these individuals appear to have had greater system understanding prior to the study's intervention. Regression analysis confirmed the ANOVA results related to treatment conditions and the lack of additional explanatory power offered by the vocabulary and mental rotation pretest. However, several demographic characteristics obtained from the self-report instrument were found to increase the explanatory power of these multivariate models albeit at small levels. In each model, the Self Regulated Learning group was

found to outperform the control group, as well as the other treatment conditions. When using the multiple choice posttest and learning gain variables, grade point average in the major was also found to explain variation. A higher GPA in the participants' selected major may indicate a proven commitment to learning and this intrinsic commitment seemed to carry over to the content area and learning episode in this study. Their academic success may also indicate the development and utilization of effective learning strategies such as coordinating multiple representations, including animations which would have benefited performance in this study. However, this variable was only found to be significant for the multiple choice posttest and not the more comprehensive mental model assessment. This difference could possibly be explained by the assessment instruments that these participants were most accustomed to completing in their prior university coursework. While many disciplines assess with multiple choice exams and written essays, most, outside of the sciences, do not include graphical representations and symbology representing target understanding of systems.

Several additional demographic variables were discovered to be significant in the ordinal regression models examining mental model posttest scores. Female participants, juniors and seniors, higher overall GPAs and more math courses were found to increase the likelihood of producing higher level mental models. While interesting and potentially informative for targeting learners who might struggle with the Hadley Cell concept, the lack of consistency for these explanatory variables across the posttest measures may indicate that the ordinal regression model is overly sensitive in identifying significant relationships. Therefore, these results should be interpreted with some skepticism and further examination in hopes of independent confirmation.

The qualitative analysis of procedural data for self regulated learning behavior produced few revelatory results beyond the recognition that learners given additional dynamic, representations chose to examine them more frequently if the opportunity existed. Since the SRL group had more opportunities to view the animations, this strategy was employed with more frequency than in the other study groups and resulted in greater content understanding given the difference noted in the mental model posttest scores. This result confirms Green and Azevedo (2007) finding that coordinating multiple representations results in a greater shift in mental model level. However, unlike Azevedo and Cromley (2004) and Hmelo-Silver and Azevedo (2006), increased frequencies of metacognitive monitoring processes were not observed in the SRL group nor the animation after group given the potential for greater prior knowledge acquisition due to the availability of the text and diagram packet. Azevedo et al. (2004) noted that metacognition was under utilized in content area novices and thus even after reading the content packet, most participants were still content area novices and did not possess this ability. Libarkin et al. (2003) noted that novice mental models are unstable and the overall low posttest scores obtained from the study's participants indicate that the target system was very complex. Transcribed participant statements included in the results section concerning the lack of isobar and animation symbology understanding poignantly illustrate this point and infer the continued instability of the participant's mental models following the learning episode.

Therefore as the prior quantitative analyses have empirically illustrated, and the qualitative analysis reinforces, the animations in and of themselves are not a standalone, quick fix to improve learners' understanding of Hadley Cell circulations. Rather, the

animations are an additional, alternate representation that if implemented at the proper time within the overall presentation and learning sequence, can improve learner understanding. However, requisite prior knowledge must exist or be obtained in concurrence with animation use to guide learner's attention to the affordances offered by the dynamic representation. These findings provide further support for prior research in the domain of learning with animations and extend this work by illuminating a sequence effect heretofore unseen.

Chapter 6. Conclusions

The findings reported in this study indicate that learners given control of the deployment of animation sequences, in concurrence with textual and diagrammatic representations of the same content, obtain a better understanding of Hadley Cell atmospheric circulations as compared to those learners without flexible control of animation viewing. Moreover, the effect of utilizing the animations as initiated by the learner while reading and examining the content materials appears to aid learners across preexisting levels of prior knowledge, verbal ability and mental rotation ability. However, given the characteristics of the sample observed in this study, most participants lacked significant exposure to the content area prior to the experiment and thus could be categorized as novices inferring that these learners benefit the most. While few additional characteristics of the sample group were consistently found to be associated with increased understanding during the experiment, self reported grade point averages in the learners' majors seemed to surface as the exception. This attribute likely did not suggest increased reading comprehension skills given the inclusion of the verbal ability measure, however, this finding may suggest that these participants were more prone to take learning activities seriously and put forth the greatest effort given their prior academic performance.

The analysis of the procedural data did not identify specific self regulated learning behaviors unique to learners given the freedom to examine the animations on demand with one exception. While the individuals selected for think aloud analysis were not observed to employ any additional types or frequencies of planning, monitoring, and task demand behaviors, a difference was noted in a single type of strategy use. Given the

opportunity to examine the additional animation representations more frequently than other treatment conditions, these participants did so resulting in a higher number of representational coordination activities.

This study's findings contribute to the understanding of learning with animation in three important ways. First, the animation sequence used in the treatment conditions were not customized for the experiment, rather they were selected from a set provided with most introductory geosciences textbooks. These animations appeared nearly a decade ago and have been adopted for use by many instructors without systematic analyses evaluating their pedagogical position and effect on student understanding. While the evaluation of textbook specific animations was not the primary goal of this research, their benefit to learners if employed in the proper temporal sequence has been confirmed for this earth science content area and sample. Prior empirical research has found that the inclusion of animations in instruction does not always guarantee increased understanding, however this study indicates that, at least for one introductory atmospheric science concept, their utilization by novice learners while reading and viewing static representations does improve system understanding. The results of this study also confirm research illuminating the frequent use of bottom up processing of animation elements by novice learners in the domain (Kriz & Hegarty 1999). The treatment condition examining the animations prior to reading the textbook material and viewing the diagrams produced lower scores on both posttest assessments than learners examining the animations after or while reading. Thus this finding may indicate that these learners were unable to selectively encode the features most relevant to system understanding given the near complete absence of prior knowledge related to the topic. This salience

effect has been noted across the interpretations of graphical representations and specific to meteorological map understanding (Lowe 2003; Mayer 1996). This effect can also be employed to explain the advantage enjoyed by novice learners in the treatment conditions in which the animations were viewed after or while reading. The addition of any system knowledge prior to observing the animations allowed these participants to approach these dynamic representations from a top-down processing perspective. Irrelevant features of the animations could potentially be ignored while attending to the most important system features. Given the constraints of the human cognitive system, the process of selectively encoding only the salient features of the animations resulted in the more efficient acquisition of system understanding.

These results must be considered in light of the study's limitations. The sample was drawn from an introductory general education laboratory science course at a mid-Atlantic regional comprehensive university. Thus most participants had not enrolled in the course given an intrinsic interest in the course's content material. Therefore cognitive effort exerted by the sample in the learning activity may have been tempered by a "requirement attitude" and may not be indicative of intrinsically interested learners (i.e., atmospheric science majors) in an upper-level course examining the same topic. An extension study including these learners should provide an interesting comparison and possible contrast to the results reported here. Similarly, procedural data was collected for only twelve students enrolled in the course. While the consistency of the reported self regulatory behaviors implies a larger sample would produce a similar distribution and frequency, this statement is only supposition without the empirical data to evaluate the claim. An increased sample size may possibly produce different self regulatory behaviors

given the larger number of participants or even given variability observed in classroom dynamics in over 15 years of teaching this content area. And as noted above, the collection of procedural data from upper class majors, theoretically no longer content area novices would serve as an interesting comparison study. One last extension idea was identified during a think aloud session when a student noted that an instructor's explanation of the animation would be beneficial to their content understanding. A direct instruction component might be merged with the animation delivery to determine whether learner understanding might increase. Animation research including animated pedagogical agents with human vocalizations to scaffold learning have been discovered to be effective therefore social interaction and discourse with an instructor, which would increase the ecological validity of this study, might be an interesting research line to pursue.

As computational power and instructional technology continues to progress, and learning modules continue to be developed and delivered in unique ways via distance learning as well as embedded within direct instruction, learners continue to face new representational methods from which to extract understanding. Textbook publishers continue to produce and provide instructors with dynamic and interactive representations across many topics that can be accessed by students to support their learning although few have been empirically evaluated. While this study found two of the animations beneficial to novice learners in a specific earth/atmospheric science domain given specific timing, continued research is necessary to determine whether animations aid in understanding additional topics. Animations continue to be produced and instructors utilizing these representations to facilitate student understanding need strong evidence to

support their inclusion given the tight temporal demands of most instructional

environments.

Appendices

Appendix A. Pressure and Wind

High pressure systems were defined to exist in locations with persistent air subsidence, either thermally or dynamically induced, where increased force is exerted on surface features. As the air reaches the surface, it diverges in all directions from the high pressure center as surface wind. Low pressure centers were defined as locations with persistent ascents, either thermally or dynamically induced, where decreased force is exerted on the surface features. Air molecules or winds converge on the surface in low pressure centers. The air traveling from high pressure centers to low pressure centers is wind whereby three forces impact the intensity and direction of these winds. These forces are the pressure gradient force, the Coriolis force and friction. Wind is the result of a pressure gradient, which is a vector quantity describing the rate of pressure change over a distance and is shown on meteorological maps by plotting isobars, lines connecting locations of equal atmospheric pressure. Air molecules move from areas of high pressure to low pressure along the pressure gradient, as gravity works to equalize the imbalance in atmospheric density, perpendicular to the isobars with the speed of movement resulting from the rate of pressure change over distance. The stronger the pressure gradient, or more rapid the pressure difference in a small geographic region, the faster the air speed or wind.

Because of the scale of atmospheric pressure systems, the distances air travels between the centers is great enough that the earth's rotation impacts wind direction when the air is plotted on a surface map. Thus, there is an apparent deflection in the direction of the wind due to the rotating frame of reference, the earth, which is termed the Coriolis

Effect. The Coriolis force is not a true force but its impact is real in modifying the direction of wind. In the Northern Hemisphere, the air is deflected at a right angle to the direction of motion while in the Southern Hemisphere; the air is deflected at a left angle to the direction of motion. The degree of deflection is a function of rotational velocity, distance traveled and wind speed. Frictional force, a force in the opposite direction of air movement, slows air speed and thus the amount of deflection as the atmosphere interacts with surface features. Thus these forces in combination with the vertical motions associated with pressure centers define circulation patterns, wind speed and directions, around high and low pressure systems in the Northern and Southern Hemisphere. High pressure systems exhibit diverging clockwise circulation in the Northern Hemisphere and counter-clockwise circulation in the Southern Hemisphere, whereas low pressure centers rotate counter-clockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere with converging winds. An understanding of both the horizontal and vertical motions of these four pressure systems was required prior to the learning episode and the understanding of following target content.

Appendix B. Ferrel and Polar Cells

A second area of thermally induced atmospheric circulation is observed over the earth's polar latitudes. In these regions, between 60° and 90° N and S latitude, very cold and dry and therefore dense air exists due to very low sun angles, long atmospheric path lengths for solar radiation allowing for increased reflection and scattering, and extended periods of little or no daylight. The resultant polar surface cools the overlying atmosphere, especially in winter, resulting in sinking air and the formation of polar high

pressure cells, called the polar highs. Although more pronounced in winter and the Southern Hemisphere given the polar landmass of Antarctica but still evident in the Northern Hemisphere over eastern Siberia, Greenland and northern Canada, and air flows from the high pressure areas toward the middle latitudes. The Coriolis force deflects these winds resulting in polar easterlies in both hemispheres. The air in the winds is very cold and very dry. Areas dominated by the polar high pressure systems and easterly winds are as dry as subtropical deserts but given their low temperatures, any precipitation that falls, nearly all in the form of snow, persists for extended periods.

Between the Tropical and Polar circulation systems lay the Middle latitudes, the region impacted most by circulation modifications due to the strong atmospheric energy gradient and local geographic features neglected by the model's simplification. In the model, the region's circulation and pressure pattern is explained as the interaction between the two adjacent thermally induced systems. Winds in the Middle latitudes result from the poleward outflow of the STH in which the Coriolis force deflects the wind to form the westerlies in both hemispheres. Where the westerly winds converge with polar easterlies, uplift and ascents occurs resulting in an area of low pressure termed the sub-polar lows, found between 50° and 60° N and S. Similar to conditions in the ITCZ, the resulting rising air leads to condensation and precipitation and therefore a pattern of humid climates. However, rather than an area of consistently warm conditions like the ITCZ, areas affected by the sub-polar lows experience oscillations between warm and cold temperatures conditions given the large scale air (or air mass) movements around the low pressure circulations. The model sometimes includes the convective circulation within the middle and polar latitudes as rising air above the sub-polar lows spills

northward and southward to the polar high and sub tropical high systems closing the tropospheric circulation loop. The Middle latitude circulation cell is termed the Ferrel Cell while the Polar cell exists in the Polar Regions.

Additional model features include the polar front and sub-tropical jet streams. The jet streams are described as discontinuous bands of high velocity upper-air winds resulting along large latitudinal temperature gradients. Two such areas are observed. The polar front jet stream is located in the area of the sub-polar lows and separates high latitude cool and cold air from middle latitude temperate conditions. The sub-tropical jet stream, weaker due to the less steep temperature gradient found in the upper atmosphere above the sub tropical high separates tropical hot air from temperate middle latitude air.

Appendix C. Pretest and Posttest Instruments

Demographic Characteristics

1. What is your gender?
　　　　a. female　　　b. male

2. What is your ethnicity?
　　　　a. white b. African-American c. Hispanic d. other _____

3. What is your age?
　　　　a. ≤ 18 b. 19　　c. 20　　d. 21　　e.>21

　　　　if other _____

4. What is you university class based on credit hours completed?
　　　　a. freshmen b. sophomore　　c. junior d. senior e. other

5. In what school is your major housed?
　　　　a. Perdue(Business) b. Fulton(Liberal Arts) c. Henson(Science) d. Seidel(Education)

　　　　4a. What is your major? _____

6. What is your current G.P.A. at Salisbury? (Estimate if unknown, freshmen leave blank)
   a. < 2.0 b. 2.0 – 2.9 c. 3.0 – 3.9 d. 4.0

7. What is your current G.P.A. in your major? (Estimate if unknown, use overall if no major declared)
   a. < 2.0 b. 2.0 – 2.9 c. 3.0 – 3.9 d. 4.0

8. How many science classes have you **completed** in college (including all campuses: Salisbury, community colleges, and other Universities)?
   a.1 b.2 c.3 d.4 e.>4

9. How many math/statistics classes have you **completed** in college (including all campuses: Salisbury, community colleges, and other Universities)?
   a.1 b.2 c.3 d.4 e.>4

10. Do you have a job?
   a. yes   b. no

11. On average, how many hours a week do you work?
   a. none b. up to 10 c. 10 to 20 d. 20 to 30 e. >30

12. Rate your general attitude toward this introductory science course on a scale of 1 to 5. (1 = dislike to 5= Like)
   a. 1 b.2 c.3   d. 4 e.5

13. Rate your general attitude toward science courses completed in college on a scale of 1 to 5.
(1 = dislike to 5= Like)
   a. 1 b.2 c.3 d.4 e.5

14. What grade do you predict that you will receive in this course?
   a. A b. B c. C d. D e. F

15. Do you use internet resources to supplement your understanding of course content?
   a. Yes b. No

16. Do you use the textbook animations to supplement your understanding of course content?
   a. Yes b. No

Appendix D. Atmospheric Circulation Post Test

Choose the correct response to each question.

 1. Global atmospheric circulation is driven by:

A) latitudinal energy imbalances
B) spring and neap tides
C) oceanic circulations
D) earth's distance from the sun
E) the moon's gravitational attraction

2. Which of the following processes and/or conditions results in semi-permanent high
   pressure systems on the earth's surface?
A) rising air
B) sinking air
C) high temperatures
D) low temperatures
E) all of the above

3. The area of consistent, low pressure located near the equator of the earth is termed:
A) the Inter-Tropical Convergence Zone
B) the Sub-Polar Low
C) the Sub-Tropical High
D) the Polar High
E) none of the above

4. Equatorial weather conditions tend to be:
A) extremely hot
B) cool
C) seasonally variable
D) consistent
E) influenced by the Coriolis force

5. A counterclockwise atmospheric circulation in the Northern Hemisphere is known as
a/an _____.
A) anticyclone
B) cyclone
C) Coriolis effect
D) pressure gradient
E) troposphere

6. Sinking air that diverges when in reaches Earth's surface is closely associated with

   _____.
A) anticyclones
B) tornadoes
C) the absence of Coriolis effect
D) the absence of friction
E) cyclones

7. Flowing air responding to the difference between higher and lower pressure is
   responding to the _____.

A) pressure gradient
B) Coriolis effect
C) anticyclone
D) intertropical convergence
E) Trade winds

8. Trade winds are found _____.
A) between 25° north and south of the equator
B) between 65° north and south of the Arctic Circle
C) centered on the longitudinal zone of the prime meridian
D) north of the monsoon regions
E) over all of the world's deserts

9. The "horse latitudes" are zones of minimal winds which are associated with the _____ system.
A) subtropical high pressure
B) trade wind
C) westerly wind
D) Polar easterly wind
E) intertropical convergence

10. Which of the following wind and pressure centers are due to the Hadley Cells?
A) jet stream
B) polar high
C) subtropical high pressure
D) westerly winds
E) sub-polar low

11. Air in a Northern Hemisphere cyclone always _____.
A) flows counterclockwise
B) flows clockwise
C) sinks
D) causes sunny skies
E) flows from the southwest

12. Which vertical air motion is found in the Intertropical Convergence Zone?
A) hail
B) snow
C) rising
D) subsiding
E) none of the above

13. Convergence is most closely associated with _____.
A) surface air in cyclones
B) surface air in anticyclones
C) cold air circulations

D) jet stream circulations
E) westerly winds

14. Which of the following is NOT descriptive of the Intertropical Convergence Zone?
A) convergence of the trade winds
B) light and variable winds
C) centered around the Equator
D) associated with rising air
E) a cloud-free environment

15. The major global wind and pressure systems
A) stay in just about the same place the entire year.
B) are controlled by Earth/Sun distance.
C) shift with the seasons.
D) are found mainly in the Northern Hemisphere.
E) seem to be independent of the jet stream.

16. In tropical atmospheric circulations, what happens to air temperature as air moves
     north and south from the equator?
A) the air temperature increases
B) the air temperature decreases
C) the air temperature remains the same

17. The reason wind exists is _____.
A) the unequal heating of the Earth system
B) Coriolis effect
C) because air is a mixture of gases
D) friction
E) altitude differences

18. This component of the global circulation is characterized by rising air, widespread cloudiness, precipitation, and migratory storms.
A) polar high
B) subtropical high
C) horse latitudes
D) intertropical convergence zone
E) trade winds

19. In tropical atmospheric circulations, surface air moves _____ the equator.
A) toward
B) away from
C) both toward and away

20. Which characteristic is associated with the Sub-tropical High?
A) cloudy skies
B) clear skies

C) very windy conditions
D) heavy rain
E) inconsistent conditions

Short Answer: Define and describe each of the atmospheric features listed below. After defining each term, describe its relationships to any of the other terms listed if they exist.

## 1. Intertropical Convergence Zone

2. Trade Winds

## 3. Anti-Trade Winds

4. Subtropical High

Essay: Describe the generalized model of Hadley Cell as discussed in your learning materials including pressure and wind features. Within your discussion, identify the locations of the major pressure centers and the dynamic and thermal mechanisms associated with their locations. Also describe the impact of these pressure systems on expected weather conditions as understood by vertical atmospheric motions. Include in the discussion the seasonal changes expected in this pattern over the course of one year. Use diagrams to describe the processes exemplified in the essay.

Diagram: Draw the generalized pattern of the Hadley Cell on the first globe below. Label the pressure centers, wind directions, and average weather conditions. On the second globe, predict the resultant pressure and winds patterns as modified by solar declination on the June solstice.

Globe 1: General pattern

30º N

0º

30º S
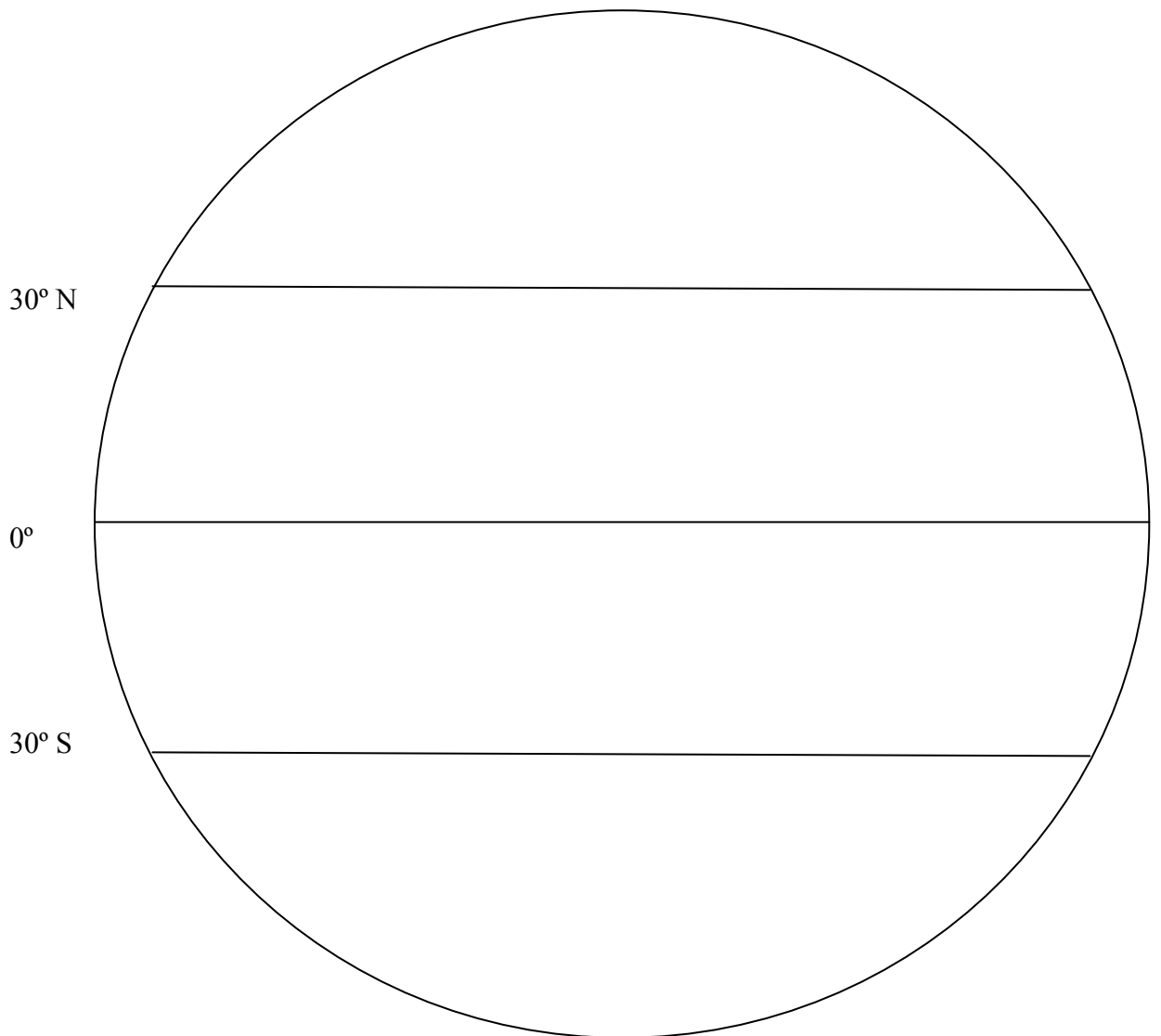
Globe 2: Pattern on June solstice

30º N

0º

30º S
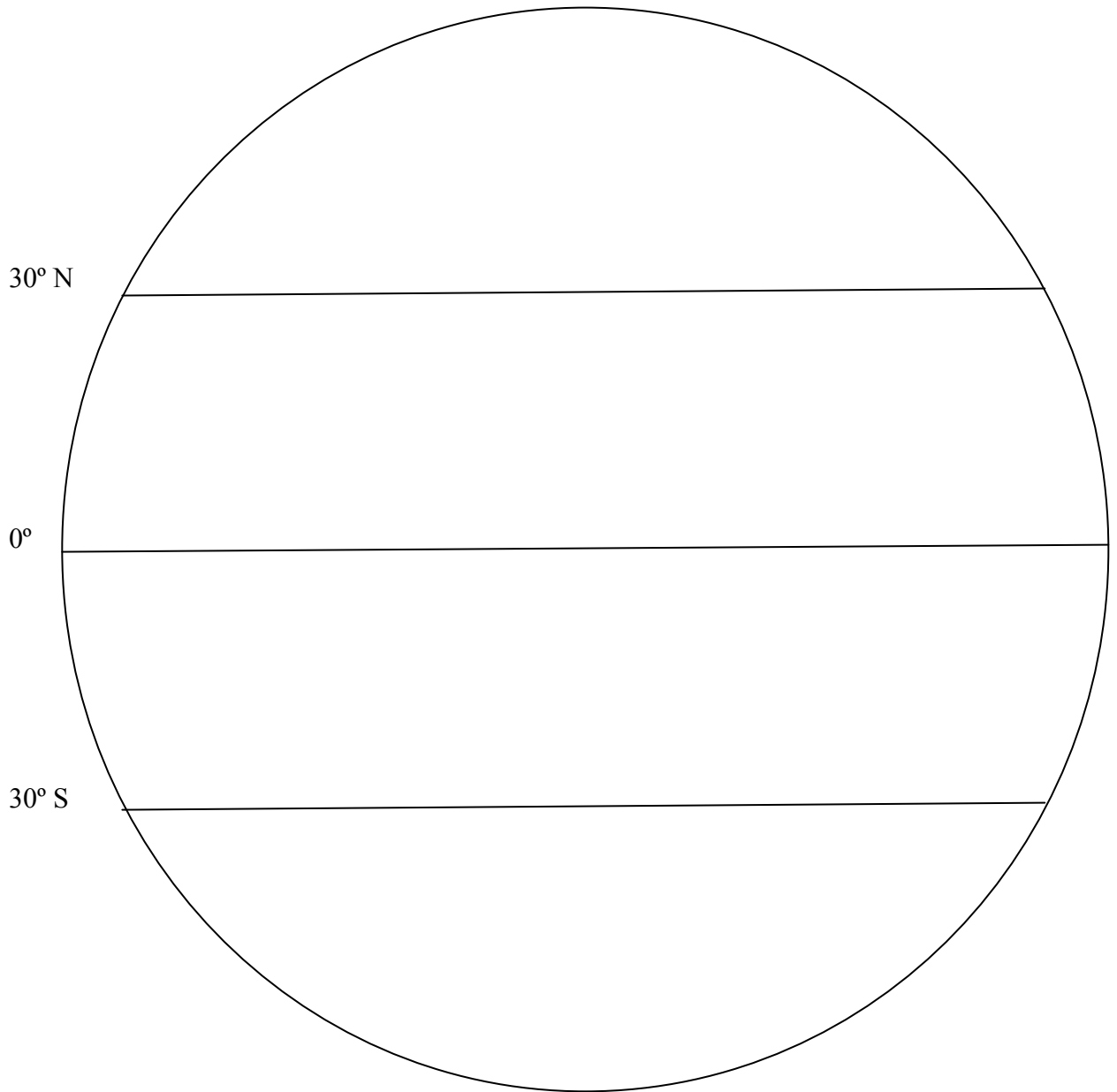
Animation: Yes/No

If yes, when did you view?   Before Reading   After Reading   While
Reading

Appendix E. SRL Behavior Codes and Explanations
From Azevedo & Cromley (2004) Coding Self Regulated Behavior

Class & Variable	Description	Example
Planning		
Planning	A plan involves coordinating the selection of operators. Its execution involves making behavior conditional on the state of the problem and hierarchy of goals and subgoals.	"First I'll look around to see the structure of environment and then I'll go to specific sections of the circulatory system."
Goals	Consist either of operations that are possible, postponed, or intended or of states that are expected to be obtained. Goals can be identified because they have no reference to already existing states.	"I'm looking for something that's going to discuss how things move through the system."
Prior knowledge activation	Searching memory for relevant prior knowledge either before beginning performance of a task or during task performance.	"It's hard for me to understand, but I vaguely remember learning about the role of blood in high school."
Recycle goal in working memory	Restating the goal (e.g., question or parts of a question) in working memory.	"Describe the location and function of the major valves in the heart."
Monitoring		
Judgment of learning	Learner becomes aware that he or she doesn't understand everything he or she reads.	"I don't know this stuff, it's difficult for me."
Feeling of knowing	Learner is aware of having read something in the past and having some understanding of it, but is not able to recall it on demand."	"Let me read this again since I'm starting to get it…"
Self-questioning	Posing a question and rereading to improve understanding of the content.	Learner spends time reading the text and then states, "What do I know from this?" and reviews the same content.
Content evaluation	Monitoring content relative to goals	"I'm reading through the info but it's not specific enough for what I'm looking for."
Identify adequacy of information	Assessing the usefulness and/or adequacy of the content (reading, watching, etc.)	"Structures of the heart…here we go…"
Monitor progress toward goals	Assessing whether previously set goal has been met.	"Those were our goals, we accomplished them."
Strategy use		
Selecting a new informational source	The selection and use of various cognitive strategies for memory, learning, reasoning, problem solving, and thinking. May include selecting a new representation, coordinating multiple representations, etc.	Learner reads about location valves, then switches to watching the video to see their location.
Coordinating informational sources	Coordinating multiple representations, e.g., drawing and notes.	"I'm going to put that (text) with the diagrams."
Read a new paragraph	The selection and use of a paragraph different from the one the student was reading.	"OK, now on to pulmonary."
Read notes	Reviewing learner's notes	"Carry blood away. Arteries-away."
Memorization	Learner tries to memorize text, diagrams, etc.	"I'm going to try to memorize this picture."
Free search	Searching the hypermedia environment without specifying a specific plan or goal.	"I'm going to the top of the page to see what is there."
Goal-directed search	Searching the hypermedia environment after specifying a specific plan or goal.	Learner types blood circulation in the search feature.
Summarization	Summarizing what was just read, inspected, or heard in the hypermedia environment.	"This says that white blood cells are involved in destroying foreign bodies."

Taking notes	Copying text from the hypermedia environment.	"I'm going to write that under heart."
Drawing	Making a drawing or diagram to assist learning.	"I'm trying to imitate the diagram as best as possible."
Rereading	Rereading or revisiting a section of the hypermedia environment.	"I'm reading this again."
Inferences	Making inferences based on what was read, seen or heard with prior knowledge.	Learner sees the diagram of the heart and states, "So the blood…through the…then goes from the atrium to the ventricle…and then…"
Hypothesizing	Asking questions that go beyond what was read, seen, heard with prior knowledge.	"I wonder why just having smooth walls in the vessels prevent blood clots from forming…I wish they explained that…"
Knowledge elaboration	Elaborating on what was just read, seen, or heard with prior knowledge	After inspecting a picture of the major valves of the heart, the learner states, "So that's how the systemic and pulmonary systems work together."
Mnemonic	Using a verbal or visual memory technique to remember content.	"Arteries – A for away."
Evaluate content as answer to goal	Statement that what was just read and/or seen meets a goal or subgoal.	Learner reads text: "So, I think that's the answer to this question."
Find location in environment	Statement about where in environment learner has been reading.	"That's where we were."

Task difficulty and demands

Time and effort planning	Attempts to intentionally control behavior.	"I'm skipping over that section since 45 minutes is too short to get into all the details."
Help-seeking behavior	Learner seeks assistance regarding either the adequateness of his or her answer or instructional behavior.	"Do you want me to give you a more detailed answer?"
Task difficulty		"This is harder than reading a book."
	Learner indicates one of the following: (1) The task is either easy or difficult, (2) the questions are either simple or difficult, or (3) using the hypermedia environment is more difficult than using a book.	
Control of context	Using features of the hypermedia environment to enhance the reading and viewing of information.	Learner double-clicks on the heart diagram to get a close-up of the structures.
Expectation of adequacy of information	Expecting that a certain type of representation will prove adequate given the current goal.	"That video will probably give me the info I need to answer this question."

Interest

Interest statement	Learner has a certain level of interest in the task or the content domain of the task.	"Interesting," "This stuff is interesting."

Bibliography

Aguardo, E. & Burt J. E. (2010). *Understanding Weather and Climate,* New York: Prentice Hall.

Anderson , J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Atkinson, R., Mayer, R., Merrill, M. (2005) Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice, *Contemporary Educational Psychology*, 30, 117–139.

Azevedo, R., Moos, D., Witherspoon, A., & Chauncey, A. (2010). Measuring the cognitive and metacognitive regulatory processes used during hypermedia learning: Theoretical, conceptual, and methodological issues. *Educational Psychologist*.

Azevedo, R., & Witherspoon, A. (2010). Self-regulated learning with MetaTutor: Advancing the science of learning with MetaCognitive tools. In M. Khine & I. Saleh (Eds.), *New science of learning: Computers, cognition, and collaboration in education*. Amsterdam: Springer.

Azevedo, R. (2009). Theoretical, methodological, and analytical challenges in the research on metacognition and self-regulation: A commentary. *Metacognition & Learning*, 4(1), 87–95.

Azevedo, R., & Witherspoon, A. M. (2009). Self-regulated learning with hypermedia. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education,* (pp. 319-339). Mahwah, NJ: Routledge.

Azevedo, R., J. T. Guthrie, & D. Seibert (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, 30(1&2) 87-111.

Azevedo, R., & J. G. Cromley (2004) Does Training on Self-Regulated Learning Facilitate Students' Learning With Hypermedia? *Journal of Educational Psychology* 2004, Vol. 96, No. 3, 523–535.

Baddeley, A. (1992, January 31). Working memory. *Science, 255*, 556-559.

Baek, Y. K. & Layne, B. H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction*, 15, 131–135.

Bodamer, D., R. Ploetzner, I. Feuerlin, & H. Spada. (2004) The active integration of information during learning with dynamic and interactive visualizations, *Learning and Instruction, 14,* 325-341.

Brewer, M. (2000). Research Design and Issues of Validity. In Reis, H. and Judd, C. (eds) *Handbook of Research Methods in Social and Personality Psychology.*Cambridge:Cambridge University Press.

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

Camp, W. G (1983). Social efficiency and vocational education: An examination of our changing philosophies. *Journal of Vocational Education Research*, 8(3), 10-19.

Carver, S. M., Lehrer R., Connell T., & Erickson J. (1992). Learning by hypermedia design: Issues of assessment and implementation. *Educational Psychologist 27*(3), 385-404.

Chandler, P. & Sweller, J. (1991) Cognitive load theory and the format of instruction. *Cognition and Instruction 8,* 293-332.

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

Clement, J. J. & Steinberg, M. S. (2002). Step-wise evolution of mental models of electric circuits: A "learning aloud" case study. *The Journal of the Learning Sciences 11*(4), 389-452.

Coll, R., & Treagust, D. F. (2003). Investigation of secondary school, undergraduate, and graduate learners' mental models of ionic bonding. *Journal of Research in Science Teaching, 40*(5), 464-486.

Danielson, E. W., Levin J., & Abrams E. (2010). *Meteorology*. New York: McGraw Hill.

Dembo, M. H. (1994). *Applying educational psychology* (5th ed.). White Plains, NY: Longman Publishing Group.

diSessa, A. A. (1993). Toward and epistemology of physics. *Cognition and Instruction 10*, p. 105-225.

Doolittle, P.E. & Camp, W.G. (1999). Constructivism: The career and technical education perspective. *Journal of Vocational and Technical Education*, 16(10).

Duschl, R.A., Schweingruber, H.A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, D.C.: The National Academies Press,

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences*. 11 (1), 105-122.

Edelson, D. C. (2001). Learning-For-Use: A Framework for the Design of Technology-Supported Inquiry Activities. *Journal of Research in Science Teaching*, 38 (3), 355-385.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, 8 (3/4), 391-450.

Edelson, D. C., Gordin, D.N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, 8(3&4), 391-450.

Edelson, D. C., & Gordin, D. (1998). Visualization for learners: A framework for adapting scientists' tools. *Computers and Geosciences*, 24 (7), 607-616.

Edelson, D., Pea, R., & Gomez, L. (1996). Constructivism in the Collaboratory. In B. G. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design*, (pp. 151-164). Englewood Cliffs, NJ: Educational Technology Publications.

Edelson, D. C. (1993). Socrates, Aesops, and the computer: Questioning and storytelling with multimedia. *Journal of Educational Multimedia and Hypermedia*, 2 (4).

Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). Kit of factor-referenced cognitive tests. Princeton, NJ7 Educational Testing Service.

Fraser, D. M. (1997). Ethical dilemmas and the practical problems for the practitioner researcher. Educational Action Research, 5(1), 161-171.

French, J., Ekstrom, R., & Price, L. (1963). Kit of reference tests for cognitive factors (revised edition) Princeton, NJ: Educational Testing Services.

Gentner, D. & Gentner D. R. (1991). Flowing water or teeming crowds :Mental models in electricity. In Gentner, D., & Stevens, A. L. (eds.). *Mental models* (pp. 99-129) Hillsdale, NJ: Lawrence Erlbaum Associates.

Goldman, S.R. (2008) Animating the issues for research in animation. In Lowe, R.K. & Schnotz, W. (eds.). (2008). *Learning with animation; Research Implications for Design*. (pp. 357-370). Cambridge: Cambridge University Press.

Goldman, S.R. (2003) Learning in complex domains: when and why do multiple representations help? *Learning and Instruction*, *13*, 239-244.

Good, T. L., Brophy, J. E. (1990). *Educational psychology: A realistic approach.* (4th ed.).White Plains, NY: Longman

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of sophisticated mental models. *Contemporary Educational Psychology*, 34, 18–29.

Greene, J. A.., & Azevedo, R. (2007). Adolescents' use of self-regulatory processes and their relation to qualitative mental model shifts while using hypermedia. *Journal of Educational Computing Research 36*(2) 125-148.

Hegarty, M. (2004). Mechanical reasoning as mental simulation. *TRENDS in Cognitive Science*, *8*, 280-285.

Hegarty, M, Kriz, S. & Cate, C. (2003) The roles of mental animations and external animations in understanding mechanical systems. *Cognition & Instruction*, 21(4), 325-360.

Hegarty, M. A. & Kozhevnikov. M (2001) A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29(5), 745-756.

Hegarty, M. Narayanan N. H. & Freitas, P. (2001) Understanding machines from multimedia and hypermedia presentations. In Otero, J. León, J. & Graesser, A. C..(EDs.), *The psychology of science text comprehension (*pp. 357-384). Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers

Hegarty, M. (2000). Capacity limits in diagrammatic reasoning. In M. Anderson, P. Cheng & V. Harslev (Eds.) *Theory and application of diagrams. Lecture Notes in Artificial Intelligence 1889*. Berlin: Springer.

Hegarty, M., Shah P.& Miyake, A. (2000) Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition*, 28(3), 376-385.

Hegarty, M. & M. Kozhevnikov (1999). Spatial abilities, working memory, and mechanical reasoning. In Gero, J. S. & Tversky B.(Eds.), *Visual and Spatial Reasoning in Design*. University of Sydney. (Internet Book Chapter-no page numbers)

Hegarty, M. and Steinhoff, K. (1997) Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual differences, 9,* 19-42.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18(5)* 1084-1102.

Hess, D. & Tassa, D. (2010) *McKnight's Physical Geography: A Landscape Appreciation*. New York: Prentice-Hall.

Hmelo-Silver, C. E. & R. Azevedo (2006) Understanding complex systems: some core challenges, *The Journal of the Learning Sciences 15*(1), 53-61.

Höffler, T. N. & Leutner, D. (2007) Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17, 722-738.

Hogan, K., Natasi, B. K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction 17(4)*, 379-432.

Holliday, W. G., Whittaker, H. G. & Loose, K. D., (1984). Differential effects of verbal aptitude and science questions on the comprehension of science concepts. *Journal of Research in Science Teaching, 31,* 877-893.

Holliday, W. G. Brunner L. L. & Donais, E. L. (1977). Differential cognitive and affective response to flow diagrams in science. *Journal of Research in Science Teaching, 14,* 129-138.

Human-Vogel, S. (2006). Students' mental models of positive mood and self-regulation in learning. *South African Journal of Psychology 36*(3) 613-633.

Jacobson, M. J. & U. Wilensky (2006) Complex systems in education: scientific and educational importance and implications for the learning sciences, *The Journal of the Learning Sciences 15(1)*, 11-34.

Johnson-Laird, P. N. (1983) *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge Mass: Harvard University Press.

Just, M.A. & Carpenter, P.A. (1992) A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99,* 122-149.

Kaplan, D. E. & Black, J. B. (2003). Mental models and computer-based scientific inquiry learning: Effects of mechanistic cues on adolescent representation and reasoning about causal systems. *Journal of Science Education and Technology, 12*(4) 483- 493.

Ke, J., Monk M., & Duschl, R. (2005). Learning introductory quantum physics: sensori-motor experiences and mental models. *International Journal of Science Education 27*(13) 1571-1594.

Kehoe, C. (1999). *The Information Processing Theory*. Retrieved Oct. 21st, 2005. http://chd.gse.gmu.edu/immersion/knowledgebase/strategies/cognitivism/informationprocessing.htm

Kozhevnikov. M. & Hegarty, M. A. (2001). Impetus beliefs as default heuristics: dissociation between explicit and implicit knowledge about motion. *Psychonomic Bulletin & Review, 8,* 439-453.

Kozma, R. (2003) The material features of multiple representations and their cognitive and social affordances for science understanding, *Learning and Instruction 13,* 205-226.

Kozma, R.B., Russel, J., Jones, T., & Marz, N. (1996). The use of multiple, linked representations to facilitate science understanding. In S. Vosniadou, E. DeCorte, R. Glaser, & H. Mandl (eds.), *International perspectives on the design of technology-supported learning environments* (pp. 41-60). Mahwah, NJ: Erlbaum.

Kozma, R.B. (1991). Learning with media. *Review of Educational Research. 61,* 179-211.

Kriz, S. & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, 65, 911-930.

Larkin, J. (1983) The role of problem representations in physics. In Gentner, D., & Stevens, A. L. (eds.). *Mental models* (pp. 75-98) Hillsdale, NJ: Lawrence Erlbaum Associates.

Lehrer, R. & Schauble L. (2003) Origins and evolution of model-based reasoning in mathematics and science. In Lesh, R. & Doerr H. M. (eds.) *Beyond Constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching.* (pp. 59-70 Erlbaum: Mahwah, NJ.

Lehrer, R. & Schauble L., (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching, 35*(1), p 3-25.

Lemke, J.J. (2001) Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching,* 38(3), 296-316.

Libarkin, J. C., Beilfuss, M. & Kurdziel, J. P. (2003). Research methodologies in science education: Mental models and cognition in education. *Journal of Geoscience Education, 51*(1) p. 121-126.

Lowe, R.K. & Schnotz, W. (eds.). (2008). Learning with animation; Research Implications for Design. Cambridge: Cambridge University Press.

Lowe, R.K. (2008). Learning from animation: Where to look, when to look. In Lowe, R.K. & Schnotz, W. (eds.). (2008). *Learning with animation; Research Implications for Design*.(pp. 49-68). Cambridge: Cambridge University Press.

Lowe, R.K. & Schnotz, W. (2005) Traitements cognitifs et fonctions , W. (2005) Traitements cognitifs et fonctions pédagogiques des animations. In J-M Boucheix, & J-F. Rouet (Eds.). *Les animations graphiques et leurs effets dans le cadre des technologies d'apprentissage*.

Lowe, R. K. (2004) Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*, 257-274.

Lowe, R.K. (2004). Animation and learning: Value for money? In R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips (Eds), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 558-561). Perth, 5-8 December. http://www.ascilite.org.au/conferences/perth04/procs/lowe-r.html

Lowe, R. K. (2003) Animation and learning: selective processing of information in dynamic graphics, *Learning and Instruction, 13*, 157-176.

Lowe, R. K. (2002) Perceptual and cognitive challenges to learning with dynamic visualizations. In Proceedings of the International Workshop on Dynamic Visualizations and Learning. Ploetzner, R. (ed.). Tübingen: Knowledge Media Research Center.

Markman, A.R. & D. Gentner (2001) Thinking. *Annual Review of Psychology*, *47,* 223-247.

Mayer, R.E., Hegarty, M., Mayer, S., & Campbell, J. (2005) When static media promote active learning: Diagrams versus animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*, 256-265.

Mayer, R., Fennell, S., Farmer, L., and Campbell, J. (2004) A Personalization Effect in Multimedia Learning: Students Learn Better When Words Are in Conversational Style Rather Than Formal Style. *Journal of Educational Psychology*. 96 (2) 389–395.

Mayer, R., Dow, G. and Mayer, S. (2003) Multimedia Learning in an Interactive Self-Explaining Environment: What Works in the Design of Agent-Based Microworlds? *Journal of Educational Psychology*. 95(4), 806–813.

Mayer, R. E., (2001). Multimedia learning. New York: Cambridge University Press.

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. (1996). Learners as information processors: legacies and limitations of Educational Psychology's Second Metaphor. *Educational Psychologist 31(3/4),* 151-161.

Merrill, M. D. (2001). A knowledge object and mental model approach to a physics lesson. *Educational Technology, 2001,* 36-47.

Merrill, M. D. (1991). Constructivism and instructional design. *Educational Technology,* May, 45-53.

Miller, G.A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Science, 7,* 141- 145.

Moreno, R. (2005) Animated pedagogical agents: How do they help students construct knowledge from interactive multimedia games? In R. Lowe & W. Schnotz (Eds.) *Learning with animations* (pp. 183-207). Cambridge: Cambridge University Press.

Moreno R. (2004) Decreasing cognitive load for novice students: effects of explanatory versus corrective feedback on discovery-based multimedia. *Instructional Science 32,* 99-113.

Moreno, R. & R. Duran (2004) Do Multiple Representations Need Explanations? The Role of Verbal Guidance and Individual Differences in Multimedia Mathematics Learning. *Journal of Educational Psychology, 96,* 3, 492-503.

Moreno, R. and Mayer R. (2004a) Personalized Messages That Promote Science Learning in Virtual Environments, *Journal of Educational Psychology.* 96 (1), 165–173.

Moreno, R. and Mayer R. (2004b) Role of Guidance, Reflection, and Interactivity in an Agent-Based Multimedia Game. *Journal of Educational Psychology.* 97 (1), 117–128.

Moreno R. (2002) Who learns best with multiple representations? Cognitive theory implications for individual differences in multimedia learning. In 14th ED-MEDIA World Conference on Educational Multimedia, Hypermedia & Telecommunications Proceedings. Denver CO June 24-29, 2002.

Moreno, R. & R. E. Mayer (1999). Cognitive Principles of Multimedia Learning: The role of modality and contiguity, *Journal of Educational Psychology, 91,* 2, 358-368.

National Research Council (2007). Taking science to school: Learning and teaching science in grades K-8. Committee on Science Learning, Kindergarten through eighth grade. R.A. Duschl, H.A. Schweingruber, & A.W. Shouse (Eds.). Washington DC: National Academies Press.

Narayanan, N. H. & Hegarty, M. (2002). Multimedia design for communications of dynamic information. *International Journal of Human-Computer Studies*, 57 (4), 279-315.

Narayanan, N. H. & Hegarty, M. (1998). On designing comprehensible hypermedia manuals. *International Journal of Human-Computer Studies*, *48*, 267-301.

Norman, D. A. (1991) Some observations on mental models. In Gentner, D., & Stevens, A. L. (Eds.). *Mental models* (pp. 7-14) Hillsdale, NJ: Lawrence Erlbaum Associates.

Park, O.-C. & Gittelman, S. S. (1992). Selective use of animation and feedback in computerbased instruction. *Educational Technology, Research, and Development*, 40, 27–38.

Perkins, D. N. (1991). Educating for insight. *Educational Leadership. October 1991*, 4-8.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.) *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic Press.

Ploetzner, R. & R. Lowe (2004) Dynamic visualizations and learning. *Learning and Instruction, 14,* 235-240.

Russell, J, Kozma, R., Jones, T., Wykoff, J. Marx, N., Davis, J. (1997). *Journal of Chemistry Education*, *74*, 330-335.

Schnotz, W. & R. Lowe (2008). A unified view of learning from animated and static graphics. In Lowe, R.K. & Schnotz, W. (eds.). (2008). *Learning with animation; Research Implications for Design*. (pp. 304-356). Cambridge: Cambridge University Press.

Schnotz, W. & T. Rasch (2008). Functions of animation in comprehension and learning. In Lowe, R.K. & Schnotz, W. (eds.). (2008). *Learning with animation; Research Implications for Design*. (pp. 92-113). Cambridge: Cambridge University Press.

Skinner, B.F. (1950). Are theories of learning necessary? *Psychological Review*, 57(4), 193-216.

Siegler, R. S. & Alibali, M. W. (2005). Information processing theories of development (pp.65-106). In *Children's thinking* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Steffe, L. P., & Gale, J. (Eds.) (1995). *Constructivism in education*. Hillsdale, NJ: Erlbaum.

Stull, A. and Mayer, R. (2007) Learning by Doing Versus Learning by Viewing: Three Experimental Comparisons of Learner-Generated Versus Author-Provided Graphic Organizers. *Journal of Educational Psychology*. 99 (4), 808–820.

Sweller, J. (1994) Cognitive Load Theory, learning difficulty, and instructional design. *Learning and Instruction 4,* 295-312.

Tarbuck, E. J. & Lutgens, F. K. (2006) Earth Science. Upper Saddle River, NJ: Pearson/Prentice Hall.

Taylor, I., Barker, M., & Jones, A. (2003). Promoting mental model building in astronomy education. *International Journal of Science Education 25(10)*, 1205-1225.

Thorndike, E. L. (1932). *The fundamentals of learning*. New York: Teachers College Press.

Tversky, B., Bauer-Morrison, J. & Bétrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies, 57*, 247-262.

Velazquez-Marcano, A., Williamson, V. M., Ashkenazi, G., Tasker, R., & Williamson, K. C. (2004). The use of video demonstrations and particulate animation in general chemistry. *Journal of Science Education and Technology, 13*(3), 315-323.

Vosniadou, S. & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*(4), 535-585.

Vygotsky, L. (1978). Interaction and learning and development. From: Mind and Society (pp. 79-91). Cambridge, MA: Harvard University Press.

Williamson, V. M. & Abraham, M. R. (1995). The effects of computer animations on the particulate mental models of college chemistry students. *Journal of Research in Science Teaching 32*(5) 521-534.

Winne, P. H., & Jamieson-Noel, D. L. (2002). Exploring students' calibration of self-reports about study tactics and achievement. *Contemporary Educational Psychology, 27*, 551-572.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 153-189). Mahwah, NJ: Erlbaum.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self- regulated engagement in learning. In D. Hacker, J.Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Hillsdale, NJ: Lawrence Erlbaum.

Winne, P. H. (1995) Inherent Details in Self-Regulated Learning. *Educational Psychologist, 30(4),* 173-187

Wirth, A. G. (1972). Education in the technological society: The vocational-liberal studies controversy in the early twentieth century. Scranton, PA: Intext Educational Publishers.

Yezierski, E. J., & Birk, J. P. (2006) Misconceptions about the particulate nature of matter: using animations to close the gender gap. *Journal of Chemical Education, 83(6)*, 954-960.

Zimmerman, B. J. (2000). Attaining Self-Regulation: a social cognitive perspective. In Boekaerts, M., Pintrich, P., & Zeodmer, M. (Eds.), *Handbook of Self-Regulation*. Academic Press.