The implications of generative learning strategies for integrating cognitive load and self-regulation theory into educational innovations

Paper presented at 24th International Congress of School Effectiveness and Improvement
Limassol, Cyprus

Gary Morrison, Professor
Instructional Design and Technology
Old Dominion University

John Nunnery, Executive Director
The Center for Educational Partnerships
and Associate Professor of Educational Research
Old Dominion University

January 2011
The typical textbook used in American classrooms, whether first grade or in college, have not significantly changed in format or design since the invention of the Gutenberg press over 500 years ago. Moreover, traditional textbooks are not interactive or adaptive to students’ needs, despite research which suggests that these features may be beneficial for learning (Lee & Park, 2008; Mayer et al., 2003; Park, 1996). In this paper we presented a research and development paradigm (AdapText Interactive) that responds to the need for improved textbook design and support for student comprehension. It expands previous research by creating larger portions of experimental text than has typically been done before, in a domain area (Earth science) that has been infrequently used in this type of research. As interest moves from printed books to electronic books, there is an opportunity to design a new generation of interactive textbooks and to move far beyond the first generation of ‘soft’ versions of the printed books. The first project we are undertaking project will take the first steps in the design of an innovative teaching technology that is based on research in instructional technology and educational psychology, to develop and test interactive i-textbook chapters that can be used as a model for the development of second generation i-textbooks.

The goals of the project arise from considering several areas of need in STEM education and textbook design. It is imperative that K-12 students have STEM education experiences which allow them to form a deep understanding of scientific principles and content. Such experiences promote the level of science literacy necessary for informed and responsible citizenship. They also allow students to begin pursuit of STEM-related career training. The focus of the project is middle school Earth science. In particular, the pedagogical resource to be developed will support students as they construct and reason with mental models of complex geologic structures and processes. Mental models are defined as simplified representations of
connected system components or causal mechanisms, that allow the learner to make predictions, simulate outcomes and solve problems (Greca 2000, Mayer, 2008; Mayer & Chandlger, 2001; Mayer & Moreno, 2003). The development of this level of understanding is critical for the individual and their role in society. For example, as a society we need students to pursue science-related careers in order to address and further understand these and other issues. At the same time, in order to participate in public discourse about issues such as global warming and the strategies that might be taken to counter it, the connections between marine and atmospheric conditions, and how global warming may impact the resources that sustain life, students need to understand Earth’s geological systems.

We believe that significant opportunities for progress towards meeting these needs exist at the middle and high school levels. From a developmental perspective, middle and high school is a critical period for building a knowledge base and deepening students’ appreciation for the scientific worldview. It is also a time when students’ cognitive development is such that for some, formal operations are achieved (Brown & Carniff, 2007). Hypothetical reasoning and abstract thinking are possible, and students can create and utilize simulations to help themselves understand relationships among system components. Thus, middle and high school students are able to develop relatively sophisticated causal models of complex phenomena. From a curricular perspective, students are expected to grasp complex ideas at a more sophisticated level than at elementary school. Students are, for example, expected to understand that “chemical pollution and sedimentation are great threats to the chemical and biological well-being of estuaries and oceans” (Virginia Earth Science Standard of Learning 11b,e). Mastering this idea requires the student to build a mental representation of factors which affect water quality and marine life and
to be able to simulate causes and consequences of changes in the chemical and biological makeup of waterways.

From a pedagogical perspective, students are expected to be able to learn effectively from informational text. While students in middle school science can and do engage with hands-on materials, textbooks remain the predominant form of instructional resource to the extent that Heinich (1984) refers to a symbiotic relationship between the teacher and the textbook that requires the teacher to translate the textbook into instruction. For students with effective reading comprehension skills and self-regulated learning strategies, informational text provides a way of deepening and consolidating curricular concepts. There are many ways, however, in which students can be deficient in their use of comprehension strategies and often times they fail to adequately comprehend informational text (Winograd & Johnston, 1982). In a science lesson, time is often devoted to supporting and promoting students’ comprehension skills. This presents a significant challenge for teachers of diverse groups of learners.

The goals of the first AdapText Interactive: i-EARTH project are to:

1. Re-design selected chapters of Earth science textbooks in order to reduce extraneous cognitive load
2. Assess learning outcomes associated with electronic versus paper Earth science textbook formats
3. Demonstrate the learning outcomes associated with interactive animation and simulation components in electronic Earth science textbook materials
4. Incorporate and assess the effect of adaptive support for students’ use of learning strategies and comprehension monitoring processes when using i-textbook materials
5. Identify and account for key learner variables that influence the effectiveness of i-textbook features such as animations, simulations, strategy prompts and feedback.

**Research and Development Design**

**Hypothesis.** The hypothesis guiding the project is that middle school students’ learning of Earth science topics can be enhanced with the use of adaptive, interactive textbook materials designed in accordance with cognitive load, multimedia learning and self-regulated learning theories. We plan to explore this hypothesis through a series of pilot studies, factorial experiments in ecologically valid settings, and focus group opportunities that ultimately will result in the development of an innovative technological resource – the interactive, adaptive, supportive middle school science i-textbook. In addition, prototypes of the materials will be tested in a controlled lab setting where physiological measures (e.g., eye tracking and electroencephalogram (EEG) data are collected to help us understand how students process text and pictures. These features are represented in the title of the project: AdapText Interactive signifies the adaptive and interactive nature of the text. i-EARTH stands for Interactive Electronic Adaptive Resources That Heighten science mastery.

**Theoretical background.** We will carefully, and with empirical data to support decision making throughout, merge cognitive load theory with principles of multimedia learning, self-regulated and generative learning theories with recent technological innovations. Through the use of a cross platform development package, we will create a platform which allows assessment of the effect of integrated graphics, animations, and simulations. However, optimizing learning in this environment also depends on consideration of the learner’s unique psychological attributes (Greene, Costa, Robertson, Pan, & Deekens, 2010). Specifically, learner processing of
multimedia components of electronically-enhanced learning materials varies according to prior knowledge, generative learning strategy use, self-regulated learning skills, and particularly spatial ability (Azevedo, Cromley & Seibert, 2004; Butcher, 2006; Grabowski, 2004; Wittrock 1974a, 1974b, 1974c, 1989, 1992; Mayer & Sims, 1994). In seeking to develop a flexible and adaptive instructional resource that facilitates the development of mental models we need to understand how learner variables interact with interactive and multimedia features of the textbook and how these influence extraneous and germane cognitive load. Before outlining the planned studies we present the theoretical background that guides the development of the interactive textbook materials, and describe the learner variables known to influence the usage and effectiveness of multimedia features.

**Cognitive Load Theory.** Prior to the past 20 years, researchers interpreted the magic number seven plus or minus 2 (Miller, 1956) as a measure of how many items a participant could hold in working memory at one time. It was not until Sweller’s (1988) subsequent research that researchers began to understand the effect of the design of the instructional materials on mental model (or schema) development as a result of limited working memory. For example, memorizing the six Noble gases and their chemical symbols is a relatively easy task. Similarly, memorizing all the metals and their symbols is a relatively easy task although it would take more time to learn than the six Noble gases. In contrast, learning to solve a typical time-distance problem is much more difficult. The basis for the difficulty lies in the content. When memorizing parts (or all) of the periodic table of elements, each element is learned independently. That is, argon is not dependent upon learning helium. The difference between these two learning tasks is described in terms of element interactivity (Schnotz & Kurschner, 2007; Sweller & Chandler, 1994). The memorization of parts of the periodic table is a task of low element interactivity,
which is to say that learning one element is not dependent upon the other. Learning to solve a
time-distance problem is characterized as having high element interactivity. A typical time-
distance problem can have 11 more elements that must be kept in working memory
simultaneously to solve the problem. Thus, this type of task can overwhelm the learner’s
working memory capacity and prevent the development of an appropriate mental model that can
be used to solve future problems.

Cognitive load is defined as the load placed on the cognitive system by a task. This load
is defined by two dimensions. The first is the load from the task which is labeled mental load and
the second is the learner’s cognitive capacity or resources, mental effort, brought to bear on the
task (Sweller, van Merriënboer, & Paas, 1998). There are three types of cognitive load. First is
intrinsic cognitive load that is determined by the level of interactivity between elements of the
task. Second is extraneous cognitive load which is the load imposed or created by the design of
the materials. For the studies, our interests focus on the management of two causes of
extraneous load. The first is the split-attention effect that occurs in most printed materials that
incorporate pictures (Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1991). The learner
typically first reads a description of one part of the illustration in the narrative and then must
locate that part of the illustration, maybe on another page, then reads the next phrase or
description, then switches to the illustration. Switching between the illustration and narrative
requires the learner to use working memory resources for administrative purposes of keeping
locations in the text and illustration and maintaining or rehearsing information during location
switching. The second cause of extraneous load is due to redundancy of information between the
text and illustration (Mayer & Moreno, 2002; Reder & Anderson, 1980, 1982). The learner could
achieve understanding from either source, but must now reconcile the differences between the illustration and narration to determine if they are equivalent.

The third type of cognitive load is germane load. Germane load refers to the working memory resources available after accounting for intrinsic and extraneous load (see Figure 1) and is used to engage in generative processing. The three types of cognitive load are additive. The example on the left in Figure 1 has overwhelmed working memory and the learner will not develop an appropriate mental model due to the lack of germane cognitive load. The example on the right of Figure 1 illustrates a reduction in extraneous cognitive load which frees working memory resources to develop a mental model, that is, increased germane cognitive load.

Research suggests that judicious application of message design, cognitive load theory, and multimedia learning principles can reduce extraneous cognitive load. Effective scaffolding and support of cognitive and metacognitive processes can maximize the benefits associated with germane cognitive load (Anglin & Morrison, 2001; Pociask & Morrison, 2008; Sweller, 1999). Optimal and efficient learning from technologically-enhanced text results from effective management of limited working memory resources (Sweller, van Merriënboer, & Paas, 1998). Research suggests that judicious application of message design, cognitive load theory, and multimedia learning principles can reduce cognitive load. Effective scaffolding and support of cognitive and metacognitive processes can maximize the benefits associated with germane cognitive load (Anglin & Morrison, 2001; Pociask & Morrison, 2008; Sweller, 1999). Therefore, the first step in our project will be to re-design textbook sections in accordance with cognitive load theory and to assess the impact of these changes on cognitive load and student performance. In particular, we will reduce the split attention and redundancy effects.
Learner Variables, Generative Strategies and the Application of Multimedia Learning Theory to i-textbook design. Over the past decade, Mayer and colleagues (e.g. Mayer et al., 2003) have conducted a number of experiments using short multimedia presentations about scientific and technical systems. These experiments have led to the articulation of multimedia learning principles, including the coherence and interactivity principles. The coherence principle clearly supports research on cognitive load. Studies have shown that individuals learn better when extraneous details are removed from informational material (Mayer, 2001, Butcher, 2006). The interactivity principle states that individuals learn better if they are able to proactively respond to and exert control over both the pace and the sequence of instruction (Mayer & Moreno, 2003). However, the results of these types of experiments present a challenge to instructional designers because conclusions cannot be equivocally drawn. Learner variables such as prior knowledge and self-regulated learning skills interact with features such as animation, complexity and interactivity to predict learning outcomes. Such individual differences have not typically been systematically addressed in studies that have been seeking to manipulate interactivity or animations.

Because individual difference variables are influential, and because middle school students require support and prompting to use generative learning strategies and execute self-
regulated learning skills (Pressley, 2000), we will assess students’ prior knowledge, generative and self-regulated learning strategies, and incorporate metacognitive support into the i-textbook pages as multimedia features are added. In spite of a wealth of research on learning from multimedia presentations, relatively few details have emerged about how students use generative learning strategies during extended, interactive multimedia learning episodes. According to Jonassen (1988), generative learning strategies “are those that require learners consciously and deliberately to relate new information to existing knowledge” (p.114). Cognitive learning strategies – techniques that support effective encoding – are included under this general heading and for our purposes include elaborative rehearsal and organizational strategies (Weinstein & Mayer, 1986). Successful strategies in multimedia learning environments and in traditional textbooks have involved prompting to use strategies such as elaborative interrogation (asking “why” questions; Smith, Holliday & Austin, 2010); the use of self-explanations during a multi-step or multi-component sequence (Chi, et al., 1989; Renkl, 2002); the use of worked examples with gradual fading of the completeness of the steps provided (Renkl, Atkinson & Große, 2000); directed re-reading of portions of text; help tools to reveal information (Renkl, 2002); summary writing; paraphrasing; and feedback regarding the degree to which a strategy was used successfully (Azevedo, 1995). When strategy prompts are integrated into the i-textbook platform and behaviors are recorded, we have the potential for students to demonstrate not only superior learning outcomes at posttest but also to demonstrate the processes that lead to successively more sophisticated levels of understanding and learning through the quality of their responses to and the need for activating these prompts.

Our approach will be to build in prompts that will facilitate strategy use. Of particular interest is the identification of when and how often these prompts should be incorporated. Some
insights may be gained from Azevedo, Greene and colleagues, (Azevedo, Moos, Greene, Winters, & Cromley, 2008; Greene et al., 2010), who documented differences in the strategic processes used by students with high or low domain knowledge. Specifically, even for students with effective self-regulated learning skills, low prior knowledge hampered learning outcomes (Greene et al., 2010). Students with both low prior knowledge and poor self-regulated learning skills may require a great deal of support. Furthermore, this support may need to be fixed rather than under the learner’s control, especially for less skilled learners.

Self-regulated learning (SRL) involves the metacognitive, motivational and behavioral control of learning (Zimmerman, 2000) and consists of three phases that are cyclical in nature. Here, we are most concerned with the metacognitive guidance of learning, which entails students’ ability to monitor and reflect upon their knowledge and the quality of their learning processes. In the Forethought stage of SRL learning goals are set, plans are formulated, and strategies are selected. In the Performance or Volitional Control stages, the learning task is embarked on through self-instruction, and comprehension and progress are monitored throughout. In the Self-Reflection performance is judged and feedback is reviewed in order to improve future learning episodes. Thus, prompting and scaffolding that is integral to the learning materials could conceivably support both the Forethought and Performance stages of SRL. It would benefit the learner who both lacks the metacognitive self-regulation to begin the SRL process, as well as supporting the learner who was aware of the need to be strategic but who lacked the specific knowledge to execute a particular strategy at an optimal point in time.

Contemporary research findings have highlighted the role played by SRL skills in accounting for variance in learning outcomes for multimedia environments (Azevedo, Guthrie & Seibert, 2004). Specifically, learning in multimedia and hypermedia settings allows the learner such great
control over navigation that metacognitive skills must be used very effectively. These findings will provide a foundation for designing a new type of interaction in the interactive textbook. We will focus our prompts in the interactive textbook during the Performance phase, in order to maximize the opportunity for students to engage in generative processing. In addition to assessing self-regulated learning skill as an independent variable, we will examine the possibility of interaction effects for different levels of SRL skill when learning from interactive i-textbook pages. We will also examine whether the expert-reversal effect can be applied from the perspective of students with high versus low SRL skills (Young, 1996). The learning outcomes of students with very good SRL skills may actually be hampered by a high degree of rigid metacognitive support. Likewise, students with very poor SRL skills may be hindered by the condition in which they have full control over their navigation and processing of the material (McManus, 2000), or they may demonstrate the ability to select the best strategy for their skill level (Ross, Morrison, & O’Dell, 1988).

**Materials and measures**

In addition to the rubric used to assess the textbook and i-textbook pages, individual measures of cognitive load, cognitive abilities, self-regulated learning, domain knowledge, domain attitudes, time on task, and process will be collected.

Cognitive Load. The most common measure of cognitive load is a subjective rating scale. Prior research has found that participants are capable of providing a numeric indicator of how they perceive their mental burden (Gopher & Braune, 1984; Paas, 1992). Moreover, deLeeuw & Mayer, (2008) presented evidence that questions about task difficulty versus effort could distinguish between extrinsic and intrinsic cognitive load. Although we will not only use self-report data, we will seek to establish whether or not it is possible to assess cognitive load in
middle and high school students using questions that parallel those asked by deLeeuw and Mayer.

The most commonly used approach to measuring cognitive load is subjective measures (Paas, Tuovinen, Tabbers, & Van Gerven, 2003; van Gog, Kester, Nievelstein, Giesbers, & Paas, 2009) consisting of multiple times or a single item. In contrast are physiological measures that can offer a continuous assessment of cognition and cognitive load. However, the use of physiological measures to assess cognitive load and cognition has the potential to a) verify subjective measures of cognitive load and b) provide important continuous indicators of cognition that can be recorded and eventually adapted as the learner interacts with the material (van Gog, Pass, & Sweller, 2010). Some measures such as heart rate and hormone levels do not appear to be as sensitive as other measures to changes in mental effort (Atonenko, Paas, Grabner, & van Gog, 2010). Other physiological measures such as eye tracking and electroencephalogram (EEG) provide continuous measures of cognition and cognitive load that are more sensitive to changes. For example, May, Kennedy, Williams, Dunlap, and Brannan (1990) reported four studies where eye tracking was successfully used to measure cognitive load. Similarly, Paas (2003) and van Gog et al. (2009) summarize additional studies that have successfully used eye tracking to measure cognitive load. The laboratory studies will use a heads-free eye tracking system and a 24-channel wireless EEG acquisition system to continuously track a learner’s visual attention while interacting with the instructional material. The eye tracker will be used to measure number and duration of fixations, along with visual search and navigation behaviors. Similar eye tracking measures have been used successfully to track cognitive processes and processing demand (Amadieu, van Gog, Paas, Ticot, & Marine, 2009; Rayner, 1998), cueing and complexity of animations (Boucheix and Lowe, 2010; de Koning, Tabbers, Rikers, & Paas,
2010), and multimedia learning (Jamet, Gavota, & Quaireau, 2008; Hyona, 2010) in prior studies.

More recently, EEG has been used to measure of cognitive load (Atonenko, et al. 2010). The use of EEGs provides researchers with a continuous rating of cognitive load that can provide a more accurate assessment of instructional interventions than a single subjective measure (Atonenko, et al. 2010). Of particular interest are the P3 measures of attention (Debener, Makeig, Delorme, & Engle, 2005). For example, various P3 measures from an EEG can provide insights into involuntary, stimulus directed attention mechanisms as well as events that cause updating of stimuli in working memory. Eye tracking will be correlated with simultaneous tracking of electrical activity in the brain using an EEG machine to measure cognition. Specifically, predominant signal peaks called event-related potential (ERP) in the P300 region will be linked to measure the presence, absence, and amount of cognition associated with each of the areas in the learning environments (Hillyard, 2009; Makeig, 2009; van Gog, Pass, & Sweller, 2010). By adding EEG and eye-tracking measures with the more traditional self-reports of cognitive load, we believe that we can obtain reliable measures of changes in mental effort. Using these data, we can then determine when we need to include prompts for either generative strategies or self-regulated learning prompts to help the learner develop a relevant mental model before the learner’s working memory is overwhelmed. For classroom-based studies, physiological measures are not feasible as they are designed for collection in a controlled lab environment. However, the use of physiological measures to assess cognitive load and cognition has the potential to a) verify subjective measures of cognitive load and b) provide important indicators of cognition. Eye tracking and EEG data, for example might reveal areas of the page that students reach before returning to a previous section, or otherwise dwell upon. These data could be triangulated with
think aloud data and may reveal an area of the text that is difficult. Subsequent iterations of the i-textbook could target this section for prompting and strategy use, such as re-reading, paraphrasing, or elaborative interrogation.

Cognitive abilities. Participants will be individually assessed on their verbal and nonverbal working memory, to provide an estimate of their working memory capacity. A modified version of Daneman and Carpenter’s listening span task (Daneman and Carpenter, 1980) will be used for verbal working memory. A modified version of the Wide Range Assessment of Memory and Learning subtest Finger Windows (Sheslow & Adams, 2004) will be used to assess nonverbal working memory. Students will be assessed on their spatial ability using a composite version of Shepard and Metzler’s mental rotation task (Shepard & Metzler, 1971).

Self-regulated learning. Self-regulated learning and generative strategy use will be measured using self-report, think aloud, and behavioral trace data. Students will take the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia & McKeachie, 1991), which is a self-report measure of self-regulated learning strategies that has been effectively modified for use in junior high and middle school as well as college level populations (Liu, 2003; Wolters, 2004). It is domain specific in the sense that individuals report on their use of generative strategies and motivational preferences for a particular course. In selected studies, students will also complete a think-aloud protocol in which they will be prompted to talk out loud about what they are doing as they are doing it. Think aloud protocols have been successfully used to capture self-regulated learning data in middle school and high school students (Azevedo et al., 2008; McManus, 2000). Finally, trace data will be gathered from the
iPad software. The usage of prompts, highlighting and note-taking, and navigational behaviors of students will be recorded. This will provide valuable feedback regarding actual usage of the interactive features of the App, and will enable triangulation with learning outcomes.

**Domain knowledge and attitudes.** A pre-test will be given to students ahead of each experimental study. The pre-test will consist of a prompt to write a short essay, followed by multiple choice questions. Posttests will include comprehension questions that draw from different levels of Bloom’s taxonomy, and will include text-explicit, text-implicit, and inference-based questions (Morrison, Ross, Gopalakrishnan, & Casey, 1995). Questions requiring simulation of a process will also be included in order to reveal the sophistication of students’ mental models for the relationships between components of specific geologic systems and processes. A short survey will also be administered to students to assess their attitudes towards learning Earth science content.

**Time on task and process measures.** Total time on task will be recorded from each iPad. This is for formative evaluation reasons as well as pedagogical considerations. In addition, this measure is of interest because research has mixed results regarding the effect of adding interactivity (Tabbers & deKoeijer, 2010) and prompting (Atkinson & Renkl, 2007) on time on task. For the paper version of the textbook pages, classroom behavioral observations will be conducted to assess time on task for a random subsample of students in each class. Process measures of usage (e.g. page turning, highlighting, clicking on graphics, and note-taking.) will be recorded from the iPad.

**Data Analysis.** Qualitative and quantitative data analytic techniques will be used as appropriate for each study. For the process data studies, think aloud data will be coded according

---

1 The term iPad is used as a generic term for any current or future pad type device that can be used in this study.
to a rubric that will establish processes and sub-processes in each phase of the self-regulated learning cycle. Students’ pre-test essays will also be coded according to a rubric to establish the sophistication of their mental model and comprehension of a particular phenomenon. Self-report survey data will be processed in order to generate total scores for particular aspects of self-regulated learning (e.g. cognitive versus metacognitive strategy use). Behavioral traces will be coded according to strategy use. Behavioral measures taken during the learning task will be recorded and in some cases (e.g. time on task) may be used as a dependent variable. For experimental studies, multivariate analysis of variance and covariance will be used (MANOVA, MANCOVA) to determine the effects of the independent variables on outcome variables such as comprehension. Additional measures and assessments will be incorporated as part of the ongoing formative and summative evaluation (see Evaluation section).

**Procedure: Sequence of studies (Figure 2).**

**Year 1.** The outcome of Year 1 will be the re-design of sections of middle school Earth science textbook chapters. Potential topics might include earthquakes and volcanoes, erosion and deposition, solar system, ocean water and ocean life, running water and ground water, and moisture, clouds and precipitation. Specific topics will be determined in consultation with teachers and partners to identify units with appropriate difficulty and that fit the classrooms’ schedules, that is, selecting topics that have not been addressed in earlier assignments. In order to achieve this goal, we will undertake three studies. In study 1a we will gather samples of 5-7 existing Earth science textbooks, randomly select pages and illustrations using a protocol and, using a rubric code features which increase extraneous cognitive load and decrease comprehension because of split-attention and redundancy effects. Graphics will be classified
according to purpose and function (Carney, & Levin, 2002; Levin, Anglin, & Carney, 1987). The primary hypothesis in study 1a is that existing middle school science textbooks contain features which contribute to split-attention and redundancy effects. A second hypothesis is that a high proportion of the images and graphics used in these textbooks are decorative, rather than representational, organizational, interpretational, or transformational in function. Levin, Anglin, & Carney (1987) reported effect sizes of .50 and higher for four of the five picture categories all categories but decorative pictures. Thus, a secondary question is whether textbooks are using pictures in a manner that supports learning (Carney & Levin, 2002).

In study 1b, N=120 middle school students in an Earth science course will review the redesigned pages in a pilot test of the revised version. This version will incorporate Integrated Graphics and no redundancy. Independent variables will include prior knowledge and self-reported self-regulated learning. Dependent variables will include self-assessments of different types of cognitive load (DeLeeuw & Mayer, 2008), time on task (Paas, Tuovinen, Tabbers, & Van Gerven, 2003), and students’ understanding of key principles via a posttest of comprehension and problem solving. Because this study is a pilot study, data will be collected with a view to making further revisions of the materials.

Study 1c will be a 2x2 factorial experiment for study 1c. With 120 middle school students, we will compare the effect of the original format textbook pages versus the Integrated Graphic textbook pages, and the effect of reading these pages on paper versus reading them electronically via an iPad². Students in the electronic reading conditions will be provided with the iPad ahead of time in order to reduce novelty effects. Students will engage in the reading task during a regular class period, complete a short distractor task, and then a posttest. Independent

² We will use the term iPad to refer to both the iPad and other similar pad type devices
measures will mirror those in study 1b. Additional dependent variables will include time on task and an assessment of cognitive load. Hypotheses are that students in both paper and electronic Integrated Graphics conditions will experience lower assessments of cognitive load, and superior performance on the comprehension posttest. We do not expect to observe performance differences for the manipulation of electronic versus paper medium. One research question of interest is if there are any differences in time-on-task with the two different media (Shepperd, Grace, & Koch, 2008). With a 2 X 2 balanced factorial design and 30 students per cell, a power of 0.82 is obtained for detecting a moderately large effect size.

**Year 2.** Year 2 begins with a study in which process data are gathered. It continues with studies which examine the effects of bringing interactive and adaptive features to the electronic learning materials, as well as the effect of including prompts for generative strategy use, and the effect of providing feedback following strategy use. Four smaller subsamples of students will participate in Study 2a, the goal of which is to gather data describing the processes used as students read the text. Eye tracking and EEG data will be collected for two groups of students as they read either the electronic original format text (n=20) or the electronic Integrated Graphics text (n=20). Think aloud data will be collected for two more groups of students as they read either the original format (n=20) or Integrated Graphics text (n=20). We hypothesize that eye tracking and EEG data will reveal evidence of lower cognitive load in the Integrated Graphics text format conditions. We hypothesize that analyses of think aloud data will reveal fewer instances of comprehension breakdown in the Integrated Graphics text condition. The data will also be analyzed to determine specific points in the text that are difficult to understand. These results will be used to identify when to add supports (prompts) in later studies.
After closely reviewing the results of study 1c and 2a, in Study 2b we will pilot test refined versions of the electronic pages. Student and teacher surveys and focus groups with 10-15 science teachers will also be conducted. The objective is to increase the richness of information about strengths and weaknesses in students’ generative learning strategies and science reasoning/critical thinking skills. We will use the resulting information to select a small number of critical generative learning strategies that will be incorporated as prompts in the Year 2 version of the materials. Prior research suggests that both pictures (Anglin, Vaez, & Cunningham, 2004; Levie & Lentz, 1982) and generative strategies (Johnsey, Morrison, & Ross, 1992; Rigney, 1978) need appropriate prompts and training in use for effectiveness. In addition, consultation with science teachers and science experts will facilitate the design of interactive elements into the textbook graphics. Interactive features will include learner-controlled labeling within the graphic.

Data from Study 2a and 2b will be used to inform the nature and position of the prompts in the final version of the learning materials for Study 2c. Study 2c will be a 2x2x2 factorial design involving (n=120) students in classroom settings. Independent variables will mirror study 1b with regard to learner differences. For the i-textbook pages, independent variables will include static versus interactive graphics, prompted strategy use versus no prompted strategy use, and feedback versus no feedback following strategy use. Feedback is expected to be beneficial. For example, Azevedo (1995) found that immediate feedback raised performance by an average effect size of .80. Static graphics will be of the Integrated Graphic type, with labels included as part of the graphic. Interactive graphics will have labels which can be viewed by touching a hot area within the image. In the prompted strategy use conditions, students will be prompted by the application to engage in a strategy prior to moving to the next section of the text. In the feedback
conditions, students will be prompted to engage in the strategy, review the product of an ‘expert’ using the strategy, and will then be prompted to compare their own work to that of the expert. Students will complete the textbook reading task during class time, will be provided with a short distractor task, and then will take a posttest. The posttest will parallel those in the previous studies but students’ notes and other physical traces of strategy use will be collected and examined. Additional measures of time on task and cognitive load will be gathered. Our hypotheses are that there will be main effects for graphic type, prompting and feedback. Interaction effects may be present, such that students who experience the interactive graphics with prompted support for strategy use and feedback, outperform students in other conditions. One research question of interest in this study is whether there is evidence of cutting and pasting for the generative strategy (Shepperd et al., 2008).

**Year 3.** Study 3a will explore the impact of incorporating animations and simulations into the textbook pages (Mayer, 2001). Animations can help the learner develop conceptual models of how the phenomena interacts (Mayer & Moreno, 2002). Similarly, simulations allow the learner to explore complex and dynamic relationships and to formulate and test hypotheses about the outcomes if parameters of a simulation were to be changed (Lee, Plass & Homer, 2006). Recent research, however, by Mayer, Hegarty, Mayer, and Campbell (2005) suggests that static pictures are more effective than animations. Animations *may* be more effective for learners with limited spatial ability (Bartholome & Bromme, 2009; Huk, 2006) or to animate processes that cannot be seen such as the movement of air or evaporation of water. This strategy can be seen as an extension of the congruence principle (Tversky, Morrison & Betrancourt, 2002), which states that effective animations and graphics must parallel the structural elements and causal behavior of the individual’s mental model during simulations. Thus, if the quality of the
learner’s mental model will be enhanced by the animation or simulation then they should be included (Reed, 2010). However, if the learner’s lack of prior knowledge is such that they would attend to irrelevant information then animations may not be the most useful multimedia strategy (Lowe, 1999). A significant advantage of the static visuals over animations is the lower cost for development. In this experiment, graphic type (interactive graphics vs animated graphics vs animation with simulation) will be crossed with explanation modality (text vs narration), resulting in a 3 x 2 factorial design. As before, we will assess the quality of prior knowledge using a single essay prompt given prior to the experimental session. These data will allow analyses to consider low and high domain knowledge as an independent variable. On the condition that prompting and feedback are beneficial to learning, all conditions will include prompting and feedback. Dependent variables will parallel those in study 2c. The hypotheses concern the improvement in learning outcome for animation versus static graphics, and the further improvement for animation plus simulation. We expect that this effect may depend on the nature of the posttest, and that students in the simulation condition may outperform other students on measures of problem solving and transfer only (Lee, Plass & Homer, 2006). We also suspect that learner variables such as prior knowledge may interact with manipulated i-textbook features.

In study 3b we pilot test materials for study 4a. A single class of students (n=25) will use the interactive, adaptive materials in order to ensure that they are sufficiently easy to use, and to inform the need for any additional prompts. A panel of teachers will provide feedback during the development of the redesigned materials.

Year 4. In study 4a we manipulate the variables of learner control and adaptive support, for three levels of self-regulated learning skill. Using self-report data, students will be classified
as having high, medium or low self-regulated learning skills. As before, prior knowledge will be assessed via a pre-test essay prompt, and will be used to determine if low and high prior knowledge interacts with the other independent variables. Self-regulated learning skill will be crossed with one of three types of text. The first will provide strategy-related prompts and feedback at the end of each section. The second will provide prompting and feedback should the student’s performance on end-of-section comprehension questions fall below a criterion level. The third type gives the learner control over whether or not to activate the prompting and feedback prior to the beginning of each section. Graphic features of the textbook will match those found to be most effective in study 3a and 3b. Dependent variables will mirror those in study 2c. We hypothesize that significant interactions will occur. Learners with low SRL skills will perform the most poorly in the learner control condition and will exhibit the best performance in the continuous support condition. Learners with high SRL skills may actually perform better in the learner control condition than in the continuous support condition. This finding has been documented as a parallel expert reversal effect with self-regulated learning as the independent variable (Kalyuga, Chandler, & Sweller, 2003; Young, 1996). Dependent variables will parallel those in studies 1c, 2b and 2c. The results of this experiment will be compared to research investigating the ability of students to control the pace and sequence of their learning (the interactivity principle), research on intelligent tutoring systems, and research on the impact of high and low domain knowledge and self-regulated learning skills on navigation through and learning from technology-enhanced text.

By the fourth year we will have gathered empirical data that will allow us to draw conclusions about the types of electronic textbook pages which best promote learning of Earth science content and the role of learner variables in predicting learning outcomes. We will
conduct a final round of focus groups and classroom observations to gather information about how i-textbook chapters could best be used in the course of a regular instructional period. We will assess students’ attitudes towards the final versions of the materials. Our final materials will include redesigned and empirically tested i-textbook chapter materials suitable for up to one marking period of an Earth Science course. During Year 4 we will produce the guidelines for publishers and authors. We will continue to disseminate the results of the project. Importantly, we will develop and give webinars to convey information about i-textbook redesign, the project’s findings, and information for teachers regarding how the materials can be used in the classroom. We will also begin to develop teacher professional development seminars and supporting materials in order to facilitate teachers’ use of the i-textbooks.

Figure 2. Timeline for the research
Conclusion

Textbooks in the United States have generally been “designed to sell,” not designed to support student learning. Even existing texts are rife with features that increase extraneous cognitive load, and make access to higher level content difficult for all students, but particularly those with special needs. Currently, e-texts do not capitalize on the computing capacity behind them to integrate well-established principles of learning theory, or to flexibly adapt to learner needs, whether the learner has executive function of self-regulation difficulties, or is an “expert.” Through this planned research and development effort, we hope to create the next generation of interactive expository materials that will increase access to advanced science content for diverse
student populations, and provide a framework for large-scale production of adaptive materials for major publishers of educational materials.
References


