Schemes for Integrating Text and Image in the Science Textbook: Effects on Comprehension and Situational Interest

Matthew O. Peterson
University of Illinois at Urbana-Champaign, USA

Science education researchers have turned their attention to the use of images in textbooks, both because pages are heavily illustrated and because visual literacy is an important aptitude for science students. Text–image integration in the textbook is described here as composition schemes in increasing degrees of integration: prose primary (PP), prose subsumed (PS), and fully integrated (FI). These schemes enable experimental research with authentic media without diminishing complexity. Seventh grade subjects used media variants in a within-subjects design (*N*=158). Comprehension tests revealed higher comprehension for FI over PP media variants in one of three sessions (*P*=0.0073). FI scores were highest for the other tests but not significantly so (*P*>0.05). A post-test situational interest measure revealed a significant preference pattern of FI>PS>PP. As textbooks follow a general PP strategy, expansion and confirmation of the results would suggest increasing the degree of text–image integration for middle school science textbooks.

Keywords: comprehension, science textbook, situational interest, text and image integration, textbook design.

INTRODUCTION

Textbooks are more than ubiquitous classroom resources for students. The textbook has long exerted direct influence on instruction as a kind of curriculum in itself (Miller & Krumhansl, 2009; National Education Goals Panel, 1998; Roseman, Stern, & Koppal, 2010; Woodward, 1993a). This influence applies in science education (Dall'Alba et al., 1993; DiGisi & Willett, 1995; Stern & Roseman, 2004), where the science textbook is a "delivered curriculum" (Chiappetta & Koballa, 2002). In science classrooms, the vast majority of both instruction and homework draw heavily on textbooks (Chambliss & Calfee, 1998; DiGiuseppe, 2014). Teacher dependence on curriculum materials in the middle school may largely be a product of so many teachers holding education rather than science degrees (Stern &

Correspondence: Matthew O. Peterson,
University of Illinois at Urbana-Champaign, 901 West Illinois Street Urbana, IL 61801
217-333-0302 USA
E-mail: mopeters@ncsu.edu

Roseman, 2004). Though textbooks exert such pervasive control, instructional goals tend to be overlooked in their development (Woodward, 1993b; Pettersson, 1998). But this situation need not be addressed with pessimism. Given the situation with textbooks in science classrooms, their improvement promises a most efficient means to, in turn, improve science education (Chambliss & Calfee, 1998). That is, textbooks, or curriculum materials more generally, represent an opportunity by virtue of posing such a problem.

In the nineties, Good (1993) found that only "one or two" articles per year in a top science education journal were devoted to curriculum materials analysis. If this number has increased, it has not done so dramatically. And a review of recent literature on textbooks in science education journals reveals a predominant emphasis on text over image—this despite the prevalence of images in science textbooks (Lee, 2010). As of 1993, almost half of the space on science textbook pages was devoted to images (Mayer, Steinhoff, Bower, & Mars, 1995). Today, some textbook publishers do not permit more than 30% of a page's space to be occupied by text (D. Decoste & F. Abd-El-Khalick, personal communication, June 24, 2015).

North American textbooks are the product of a two-stage process (DiGiuseppe, 2014). A development team including authors, editors and marketing managers determines the curriculum and contents as the "principal decision makers" (p. 1064). The content is then handed off to a production team including designers, illustrators, and photographers. It is important to note that under this production model, designers are not involved until after content has been developed, and after the principal decision makers have done their job. It is perhaps then of little surprise that images in textbooks are often "content-supporting" rather than "content-extending" (Woodward, 1993b). Worse yet, images in textbooks can be "content-irrelevant," or decorative (Levin, 1979; Pettersson, 1998; Woodward, 1993b).

It has long been established that visual representations aid in the understanding of textual content (Levie & Lentz, 1982; Tversky, 2011; Van Genuchten, Scheiter, & Schüler, 2012; Vekiri, 2002), including in the context of science curriculum materials (Finson & Pederson, 2011; Lee, 2010; Mayer & Gallini, 1990). While informational materials constructed for experimental studies may isolate text–image constructs (such as captioned images), authentically illustrated media like the science textbook are immersive spaces with multiple layers of text and image at varying levels of integration. The importance and subtlety of text–image integration has not escaped researchers in science education (Dall'Alba et al., 1993; Dechsri, Jones, & Heikkinen, 1997; Mayer et al., 1995; Mason, Tornatora, & Pluchino, 2013), though work focusing specifically on integration as independent variable is rather rare. Mayer, Steinhoff, Bower, and Mars (1995) provided a "generative theory of textbook design" that is predicated on the integration of text and image, for which they tested aspects of integration with informational material on lightning. Mayer et al. summarized their theory: "The overarching theme of a generative theory of textbook design is that features of the text design affect the degree to which readers engage in cognitive processes necessary for meaningful learning" (p. 32). Their findings suggest that integrated media for science is superior to cases where text and illustration are presented in relative isolation. Mayer et al. focused on two aspects of integration: proximity (how close related text and image are to each other) and annotation (where images are clarified with descriptive text). This approach provided a beginning for a prescription for the design of illustrated informational media. Vekiri (2002) provided an extensive literature review that indirectly builds on this basic model, documenting many studies on "graphical displays" aimed at learning. While the literature is far from providing for a consensus on details, such as learner characteristic effects (for whom and when a kind of visual representation is effective), there is copious and variegated support for text–image integration.

The present paper examines the impact of three composition schemes of text–image integration in a science textbook on selected cognitive and affective learner outcomes. The experimental study described here, which was conducted in the United States, departed from the aforementioned textbook production model, which delays designer involvement until content has been codified in a central text. Instead, the experimental conditions were created by producing three levels of text–image integration informed by theoretically-grounded design principles (see: Mayer et al., 1995; and Vekiri, 2002). This paper ultimately aims to inform the improvements to the science textbook as a means to further science education. The impact of integration, of course, could come in many forms. In considering the textbook as an experiential artifact—something that is used rather than simply digested—three major potential areas of impact, in sequential terms as part of a textbook experience, are: (a) the reader's initial impression of the material, (b) his/her active engagement with the material, and (c) the lasting impact of this engagement. The latter stage was deemed beyond the scope of the study detailed herein, given limited time for access to the subjects, 7th grade students. The main focus was with the former two stages, which are here related with interest and comprehension, respectively. (In this case comprehension refers to the reader's ability to apprehend information and explore concepts during media use.) The additional focus on the relationships of "textbook characteristics" to student interest was based on Chiappetta, Fillman, and Sethna's (1991) call for science education researchers to go beyond comprehension when examining impacts of science textbooks. Thus, this study aimed to answer paired questions: How do different schemes for text–image integration in a science textbook impact middle school students' (1) comprehension of science content, and (2) interest level in curriculum materials?

Conceptual framework

Schemes of text–image integration

Pictorial representations are important in science education, both because images are ubiquitous in textbooks and because visual literacy is a necessary skill in science (Lowe, 2000). Tversky (2001) characterized diagrams and pictures as externalized thoughts. According to cognitive load theory (Plass, Moreno, & Brünken, 2010), dual coding theory (Sadoski & Paivio, 2001), and working memory theory (Baddeley, 1998), concepts are mentally represented in either verbal or visual codes, meaning that these distinctions exist not just on the page as information formats, but during comprehension processes and in long-term memory alike. Thus diagrams and pictures externalize thoughts that are inherently visual, through and through.

Mayer et al. (1995) established the importance of text and image in the science textbook with their generative theory of textbook design (in their study the content focus was lightning). Mayer et al. outlined the process a learner goes through as he or she mentally integrates information presented in separate codes as text and illustration (**Hata! Başvuru kaynağı bulunamadı.**). The learner first must select relevant information from text and illustrations. Because this part of the process is selective, it can of course be done well or poorly. Selected information becomes a text or image base in short-term memory. This information must then be organized into referential connections, which will persist as integrated in long-term memory.
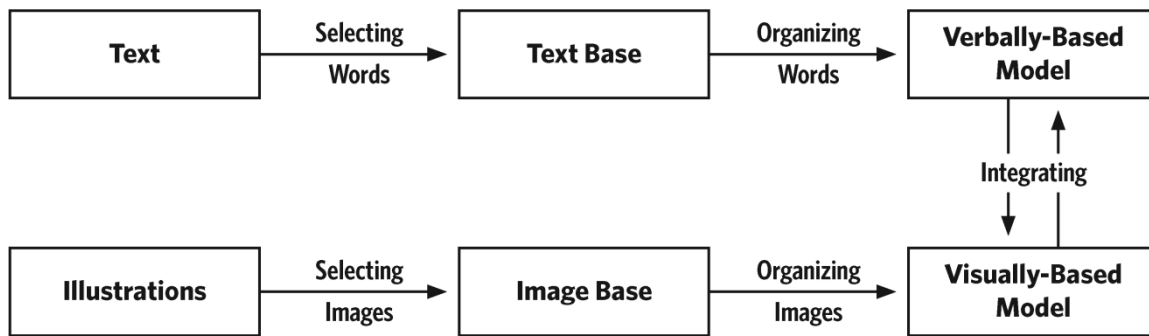
Figure 1. Mentally Integrating Text and Illustrations (redrawn from Mayer et al., 1995).

What is especially important here is that representations on the page retain their code-specific properties throughout the process. This is not to claim a perfect one-to-one correspondence between illustration and its generated mental imagery. But illustrations are encoded distinctly from text: text and image are not mentally jumbled together into some new singular format.

Mayer et al.'s generative theory of textbook design is predicated on the importance of the learner building referential connections between pieces of information while both verbal and visual representations are held together in short-term memory. This is best accomplished by integrated text and illustration on the page. Mayer et al.'s results supported this assertion. Their findings on integration are consistent with psychological principles of design (see esp. Vekiri, 2002, for a review). When Mayer et al. attempted to separate out two aspects of integration, proximity of image to corresponding text and annotation (or non-annotation) of images, they were unable find an effect of proximity alone. However, as in many cases of experimental study on text–image integration, their stimuli were not fully formed authentic informational media.

Inherent to the present study was an emphasis on fully formed authentic informational media, as complete science textbook pages. When dealing with authentic illustrated media, distinctions between annotation and text–image proximity can quickly dissolve. In Mayer et al.'s annotated–proximate stimulus materials, a dominant prose passage was displayed next to an annotated image. The caption below the image is easily identified as annotation. And proximity is measured in distance from text passage to image. But within an extensive and complex diagram, where individual pictures blend into one another and text is distributed in chunks (not in a prose passage), what counts as a caption or annotation exactly? The division of annotation and proximity is predicated on a text dominant design strategy. Schemes that capture aspects of both of these models of text–image integration are proposed by the author and outlined here. It is these schemes that are investigated in this study.

A *prose primary* (PP) scheme features a central prose column or columns (examples of PP and the following schemes are presented in **Hata! Başvuru kaynağı bulunamadı.**). The strict grammatical rules on which language is dependent oblige a specific reading order. It is difficult to enter a paragraph halfway through, or to scan prose for specific information without consenting to a line-by-line reading. Prose primary is thus inherently a serial system of organization. When images are included, they are marginal to the text, and the reader must determine when to break the prose sequence to attend to them. Figure references embedded in

the prose can suggest such breaks. Most American textbook pages take the form of PP layouts.

The most extreme form of text–image integration is a *fully integrated* (FI) scheme. Diagrams themselves are fully integrated. An entire spread in a book can be fully integrated. FI layouts do not include prose, though text is present in more discrete "chunks": as labels, phrases, or isolated sentences. Where a PP layout is serial in structure, an FI layout is parallel in structure: there is neither a singular entry point nor a proper reading order. The reader determines both individually by virtue of reading. It is not immediately apparent which integration extreme presents a more difficult reading experience, or for which type of reader. PP requires minimal decision-making in terms of path, but reading in the traditional sense (that is, processing text) is far from simple. FI requires lower traditional literacy, arguably, but counters that with a potentially overwhelming set of options. In Mayer et al.'s terms, FI would appear to make selecting words and images more of a challenge, requiring more exacting determination of the relevance of pieces of information. When readers are faced with illustrated media, the first couple fixations serve to process the "gist of the scene" before more detailed study begins (Carroll, Young, & Guertin, 1992). As long as the relationships in an FI layout are meaningful, there is a "scene" or scenes to be processed, which provides some grounding for the reader's determination of path. PP layout, in contrast, is dependent upon learned rules and standards, and doesn't appear to present a coherent scene from which readers can make task-dependent determinations.

A third, intermediate degree of text–image integration is included to provide graded differentiation and a theoretical middle ground: the *prose subsumed* (PS) scheme. Prose subsumed layouts include sections of prose, with a lesser sense that those sections must be read in a particular order, like a series of captioned images, where the captions are brief text passages. Western readers will likely address sections in a left-to-right pattern when reading for completion. However, readers can also scan or browse sections when using media as a reference. The hierarchical positioning of images in a PS layout reinforces the sense that prose sections can be ordered at the reader's discretion. Each section of prose, in addition to a title, appears below an image. Thus the relationship of image to text is largely flipped from PP to PS: images are marginal in PP; in PS the text supports the hierarchically "higher" images. Prose subsumed layouts are essentially sets of captioned images. As in PP, the "images" can at times take the form of diagrams, themselves fully integrated but isolated zones within a broader layout.

These three schemes represent increasing levels of text–image integration: PP to PS to FI. The North American textbook production model inherently favors the PP scheme by virtue of excluding designers from the development of content as prose is generated.

The suitability of a text–image integration scheme must depend upon the information represented. More integrated schemes should be of interest in science education given the importance of visual literacy in science (Akaygun & Jones, 2014; Ainsworth, 1999; Lowe, 2000). But it seems likely both audience (from grade school to higher education learners) and subject area (from cell theory to Earth science) determines scheme relevance at any given time.

The present study seeks to connect text–image integration scheme is science curriculum materials to learner comprehension of information and interest in media. The constructs of comprehension and situational interest are briefly discussed before the experiment is outlined.

Figure 2. Three Text–Image Integration Schemes. [Adapted from MCDOUGAL LITTELL SCIENCE, North Carolina Edition, Student Edition, Course 3, by Trefil, Calvo, & Cutler. Copyright © 2005 by McDougal Littell. All rights reserved. Adapted and reprinted by permission of the publisher, Houghton Mifflin Harcourt Publishing Company. Any further use is strictly prohibited unless written permission is obtained from Houghton Mifflin Harcourt Publishing Company.]

Comprehension

Comprehension of media, inclusive of text, image, and schematic arrangement, is described in many reading comprehension models (Mason, Tornatora, & Pluchina, 2013). The models agree on a general conceptualization of reading comprehension as a constructive process wherein the reader activates background knowledge and integrates it with the information he or she is reading, into coherent mental representations. Comprehension, as presently conceived and consistent with the New Taxonomy of Educational Objectives (Marzano & Kendall, 2007), involves integrating and symbolizing. To integrate, a reader must identify basic structures of meaning and be able to identify elements as critical or noncritical. To symbolize, a reader must construct a useful and accurate symbolic representation of information encountered, and do so according to how critical individual pieces of information are to the task at hand. This view is complementary to Mayer et al.'s (1995) conceptual framework of selecting, organizing, and integrating information during reading tasks, with both text and image. Comprehension of pages in a science textbook is a question of the present, of reading and referencing. According to cognitive load theory (Plass, Moreno, & Brünken, 2010; Schnotz & Kürschner, 2007), which considers human limitations in learning, it is possible to understand something without learning from it, though this is rare. Understanding in cognitive load theory, consistent with comprehension here, refers to a learner's ability to hold all task-relevant pieces of information in working memory while completing a task.

Situational interest

The textbook finds itself in a peculiar situation amongst books: its readership is largely composed of unwilling readers. This seems to suggest that engendering reader interest is particularly valuable. Interest provides "incentive to selectively attend and process information when motivation [is] low" (Schraw & Lehman, 2011, p. 27; after Asher, 1980). Interest has been found to increase attention and learning. Interest can be personal or situational (Schraw & Lehman, 2011). Personal interest is especially related to prior knowledge and changes only slowly over time. Situational interest is environmentally and spontaneously activated, context-specific, and influenced by prior knowledge but to a lesser degree than personal interest. Thus situational interest is more readily manipulated.

Interest is an important motivational variable in media-based learning. Interest encourages readers to go deeper into a text (or image) to construct meaning, and can even help learners transcend reading level to a degree (Schiefele, 1991; Worthy, 1996). Increased interest is directly related to increased reading comprehension and learning (Alexander, Jetton, & Kulikowich, 1995; Hidi, 2006), though this is not universally accepted (Rotgans & Schmidt, 2014).

Interest has been identified as an important factor in science education and scientific literacy (Krapp & Prenzel, 2011; Swarat, Ortony, & Revelle, 2012). It is currently receiving increased attention in science education journals (Lin, Hong, & Chen, 2013, noted a recent increase in articles on interest in a major journal).

Interest in media, as defined for this study, is a form of situational interest. The presence of images and their integration with text represents an extension of text-based interest studies (Krapp, Hidi, & Renninger, 1992; Mitchell, 1993; Schraw, 1997), which have tended to focus on editorial issues internal to prose.

Research design

Research design overview

This within-subjects experiment measured comprehension of science curriculum materials (textbook pages) through what were essentially open-book tests conducted in the subjects' regular classrooms. The textbook pages took the form of prose primary (PP), prose subsumed (PS), or fully integrated (FI) media. Following the comprehension tests, a post-test measured situational interest by asking subjects to select most and least interesting examples of textbook pages that varied according to text–image integration scheme: PP, PS, and FI.

School and sample

The subjects were 158 7th grade students enrolled in a middle school in the southeastern United States. The participating public magnet school was located within a city with a population in excess of 400,000. At the time of the study, the school had roughly double the population of African-American students than the national average, but with Hispanic/Latino students 2% below the national average. Free and Reduced Lunch (FRL) enrollment was 36.3%, lower than the national and state averages of 44.6% and 45.9%, respectively. FRL is reported here as a general indicator of economic status that is readily available in school statistics.

All 7th grade students at the participating school were invited to participate in the study. All sessions occurred during the subjects' regular science classes at their assigned seats, to maximize ecological validity. Subjects were distributed across 8 total science classes between two teachers (the teachers oversaw 4 classes apiece). One teacher had desks arranged in rows facing forward, while the other teacher utilized pods of roughly 4 desks apiece. Both teachers granted access to their full science periods over 4 successive Mondays.

Independent variable and stimulus materials

The independent variable of the study is text–image integration scheme, as expressed in textbook page layout and content. Stimulus materials were prepared for the study by the author, who was a practicing graphic designer before entering into doctoral studies. These media took the form of single spreads—left- and right-hand facing pages together—from a science textbook. The prose primary (PP) media variants were adapted from existing eighth grade textbook pages (Trefil et al., 2005). An 8th grade textbook was used to ensure that the content was unfamiliar to the subjects, who were in the 7th grade. The participating teachers confirmed that the topics were not familiar to their students. Prose subsumed (PS) and fully integrated (FI) media variants were created from the existing prose primary versions. The media conversion process is explained in some detail later in the paper. One set of conversion examples is shown in thumbnail in **Hata! Başvuru kaynağı bulunamadı.**.

Three sets of media variants were created for the study, each with unique content according to the original PP spreads: (1) divides and drainage basins, (2) lakes and lake turnover, and (3) fossil fuels and coal.

Dependent variables and their measurement

The dependent variables are (1) comprehension of concepts in text and image formats, and (2) situational interest in science curriculum materials, henceforth *comprehension* and *media interest*.

Subjects participated in three comprehension tests and a post-test. Each session covered one content lesson such that all subjects were engaging with the same subject matter at the same time (e.g. "lakes and lake turnover"). Subjects were instructed to read their given media variants once before they began comprehension testing. Media variants were printed and folded to simulate a book's spread. Subjects completed selected response comprehension items using online survey software (SurveyGizmo) on individual laptop computers. Subjects used the media variants during testing to separate any memory effects from comprehension. Reading and testing duration varied greatly per subject, with most subjects finishing in well under 20 minutes. Each subject did his or her work individually.

The participating teachers were in the room while subjects completed the tests. Each session was conducted by the author or a research assistant, as the two participating teachers' classes ran simultaneously.

The post-test addressed media interest, the second dependent variable. Immediately after completing the third and final comprehension test, the survey software linked to a separate survey page. Subjects were shown thumbnail views of each set of media variants, broken down into individual pages, and clicked on the media variant they found to be "most interesting." Subjects were thus comparing, for instance, all three left-hand pages of the "lakes & lake turnover" media variants. They subsequently made selections of "least interesting" for the same sets. Both items and thumbnails were randomly ordered per subject.

Sessions occurred one week apart from each other. A science self-efficacy inventory was administered in the week preceding the first comprehension test. Self-efficacy, or "confidence in one's abilities to successfully complete tasks," has been shown to contribute to academic performance (Lawson, Banks, & Logvin, 2007, p. 707). The science self-efficacy instrument used was developed by Nietfeld, Cao, and Osborne (2006), for whom it proved an internally reliable survey predictive of academic performance. This measure was used here as an adjunct for academic performance because there was no practical means by which to acquire proficiency scores (from either standardized tests or previous class performance). Science self-efficacy ratings provided a means to interrogate between-group differences that was separate from comprehension results.

Group assignment. Subjects were assigned to one of six sequence groups according to a "Williams design" 6×3 modification to a 3×3 Latin square, in a within-subjects design. The Williams design controlled for order effects: half of the PP subjects from session #1 would move on to PS the following session, while the remainder received FI.

It was assumed that classes were not equivalent. Thus students were randomly assigned to one of the six sequence groups in balanced numbers in each class.

ANOVA tests failed to find differences in the sequence groups considering either total comprehension test scores (P=0.3932) or science self-efficacy rating (P=0.2223), which indicates self confidence in science study (Nietfeld et al., 2006).

Media variant development

As this study utilizes holistic alternative versions of two-page spreads, media variant equivalency is a complicated matter. The author, previously a professional graphic designer, created the media variant sets used in the study.

Media variant development began with the original prose primary (PP) pages. Pages, with their corresponding topics, were selected according to "picturability": enough information in the original text format had to be efficiently reproducible in image format. The underlying conceptual information in the selected pages had to be extensive enough for testing subjects across 12 unique test items. Alternate versions of the original PP pages were created using fonts, type size, leading, and

spacing considerations that would be consistent through all generated media variants. Typographic decisions were guided by the original design, such that the PP variants were as similar as possible to the originals. The session #2 PP media variant is presented in **Hata! Başvuru kaynağı bulunamadı.**.

Figure 3. Comprehension Test #2 "Lakes & Lake Turnover" Prose Primary (PP) Media Variant. [Adapted from MCDOUGAL LITTELL SCIENCE, North Carolina Edition, Student Edition, Course 3, by Trefil, Calvo, & Cutler. Copyright © 2005 by McDougal Littell. All rights reserved. Adapted and reprinted by permission of the publisher, Houghton Mifflin Harcourt Publishing Company. Any further use is strictly prohibited unless written permission is obtained from Houghton Mifflin Harcourt Publishing Company.]

The prose subsumed (PS) media variants were designed second (see **Hata! Başvuru kaynağı bulunamadı.**). Prose was broken down into roughly four discrete sections. Each section, according to the PS strategy, had to be anchored with an image as an entry point, "higher" in the hierarchical structure than the sections of prose themselves. In some cases these images were in the original PP pages and were merely rearranged and resized. In other cases, new images were included. Some textual information was converted into images, but this was minimal in the context of a PS layout. When possible, the text was edited to be shorter and more immediate per prose section. Good faith attempts were made to ensure the quality of the PS variants, such that while being consistent with the outlined strategy, each was expertly made as a best possible version of itself.

The fully integrated (FI) media variants took more time to create than their PS counterparts (see **Hata! Başvuru kaynağı bulunamadı.**). Additional images were consistently found or generated. Images had to create spaces in which text would be placed. Text had to be broken down to small "chunks" and divorced from its previous prose structure.[1] Word count subtly decreased from PP to PS (e.g. 633 to 624 for "lakes & lake turnover"), and much more dramatically from PS to FI (e.g. 624 to 435 for the same).

Figure 4. Comprehension Test #2 "Lakes & Lake Turnover" Prose Subsumed (PS) Media Variant. [Adapted from MCDOUGAL LITTELL SCIENCE, North Carolina Edition, Student Edition, Course 3, by Trefil, Calvo, & Cutler. Copyright © 2005 by McDougal Littell. All rights reserved. Adapted and reprinted by permission of the publisher, Houghton Mifflin Harcourt Publishing Company. Any further use is strictly prohibited unless written permission is obtained from Houghton Mifflin Harcourt Publishing Company.]

Figure 5. Comprehension Test #2 "Lakes & Lake Turnover" Fully Integrated (FI) Media Variant. [Adapted from MCDOUGAL LITTELL SCIENCE, North Carolina Edition, Student Edition, Course 3, by Trefil, Calvo, & Cutler. Copyright © 2005 by McDougal Littell. All rights reserved. Adapted and reprinted by permission of the publisher, Houghton Mifflin Harcourt Publishing Company. Any further use is strictly prohibited unless written permission is obtained from Houghton Mifflin Harcourt Publishing Company.]

The conversion process was generally consistent with myriad psychological principles of design as reported by Vekiri (2002). But design decisions were primarily guided by the spatial contiguity principle (ibid.), which states that related pieces of information should be placed near each other. Repeated attempts to convert as much text to image as possible also guided conversions. It should be noted that FI media variants were, more than anything, schematic representations of information, akin to extended diagrams.

DATA ANALYSIS

Comprehension tests

Comprehension data collection

Subjects were instructed to read the media variant they received and then turn to the comprehension test in the online survey instrument. The media then became a resource. The comprehension test consisted of 12 items for each session. Each item was in selected response format with four options. Comprehension test items were divided into 6 factual items and 6 conceptual items, according to the classification of knowledge in Bloom's Revised Taxonomy (Anderson et al., 2001). Factual knowledge covers terminology and specific details and elements. It has a low level of abstraction. Conceptual knowledge, more complex and organized, covers classifications and categories, principles and generalizations, and theories, models and structures. For the comprehension test items, those that could be answered by finding the right piece of information at the critical locus in the media were considered to be factual. Those that required the subject to piece together information found at multiple loci, or otherwise required inference, were considered to be conceptual. Both factual and conceptual items were included to ensure that test items covered comprehension more broadly and best mimicked the variation in science testing.[2]

The following item covers factual knowledge. The key is indicated with an asterisk. The proportion of subjects selecting each option is in brackets.

- Which of the following is NOT a way that lakes can form?
- o Huge sheets of ice scrape out depressions. [0.01]
- o Igneous rock is pushed up by geologic forces. [0.84]*
- o Inactive volcanoes leave craters behind, which collect water. [0.08]
- o Water collects in low parts of land. [0.06]

The following item covers conceptual knowledge according to the study's distinction.

- If you were going on a fishing trip in a northern state, why would you want to know about lake turnover?
- o So I can be sure that all of a lake's nutrients are not floating on the surface. [0.17]
- o So I know if any fish are alive at the time. [0.08]
- o So I know at what depth the fish in a lake might be found. [0.71]*
- o So I could determine if a body of water is a lake or a pond. [0.04]

The online survey instrument presented subjects with all 6 factual items on one screen. After the factual items were completed, the subject could click to the next screen, which contained all 6 conceptual items. The order of the items was randomized within each page per subject, as were the items' options. Subjects clicked a button to the left of an option to select it.

Comprehension item development

Some of the test items came from the same textbook as the original prose primary variants, while the author developed the remainder. The participating teachers were consulted to confirm that the subjects had not previously engaged with the material. For each of an item's distracters there was some realistic misconception accountable for selection. Furthermore, the media variants were consulted to ensure that each distracter could serve its purpose across the set. Likewise, each item had to be solvable using all media variants. Following the advice of Reynolds, Livingston, and Willson (2006), the author constructed comprehension items and then returned to them after a break of at least a full day. This made it possible to see the items with "new eyes," which quite often led to alterations to the language of the items. Critical information for each test item, as related to the stem, key, and distracters, was located in and mapped onto all media variants with notation. This determined specifically how each item could be solved, and helped determine which items were conceptual and which were factual. It also helped ensure that no zone of any media variant was disproportionately critical or irrelevant. Care was taken to ensure that neither keys nor distracters tended to be longer or shorter across items.

Comprehension item analysis

Following comprehension tests, item analysis was conducted according to strict protocols. Items were interrogated by up to four methods: (1) item difficulty index, (2) item discrimination index, (3) item-total correlation coefficient (a second measure of discrimination), and for items flagged by any of the previous methods, (4) distracter analysis. These methods were conducted according to standards set out by Reynolds et al. (2006). For a test item to survive item analysis it had to pass certain score thresholds for the first three measures, ensuring a good rating. Item analysis resulted in one item apiece being removed from the comprehension tests, resulting in a final retained 11 items apiece.

Comprehension data analysis and reliability

Following the removal of items during item analysis, reliability for each 11-item comprehension test was calculated using Cronbach's alpha. Comprehension test #1 rated as *acceptable* (according to Reynolds et al. [2006]; $\alpha=0.6234$, $N=146$). Comprehension tests #2 and #3 rated as *good* ($\alpha=0.7821$, $N=143$; and $\alpha=0.7156$, $N=148$; respectively).

ANOVA tests were conducted for each comprehension test against the null hypothesis that text–image integration scheme has no impact on comprehension. In the case of rejection of the null hypothesis ($P\leq0.05$), a Student's *t*-test was conducted to isolate the significant differences among the pairings of PP/PS, PP/FI, and PS/FI.

Comprehension results

Comprehension test score means are reported in **Hata! Başvuru kaynağı bulunamadı.**. Mean scores indicate the number of items solved within the sets of 11 items. Comprehension test #1 had an aggregate mean score of 5.9178, considerably lower than #2 and #3, at 7.8741 and 7.2500, respectively. In each case mean scores for FF media variants were the highest, though the differences are subtle in the tests #1 and #3. The PP variant mean scores were lowest in the first two tests.

Subjects in comprehension test #2 solved 1.5 more items on average with the FI variant than the PP variant. An ANOVA test found significant differences in that test ($P=0.0202$, $N=143$). A Student's *t*-test revealed that FI scores were significantly higher than PP scores (FI>PP, $P=0.0073$). PS scores only came close to achieving significance over PP scores ($P=0.0557$). ANOVA testing found no significant

difference in the mean scores for tests #1 or #3 (P=0.9332, N=146; P=0.4937, N=148; respectively). The relatively "poor" results from test #1 are possibly the result of the dramatically lower mean scores. Difficulty levels for performance tests too far from an ideal fail to differentiate performance levels (Reynolds et al., 2006).

Thus there are mixed results for research question #1: How do different schemes for text–image integration in a science textbook impact middle school students' comprehension of science content? There are indications that FI may better promote comprehension than PP, but such indications are tentative given only one of three tests with significant results.

Information was controlled across media variants. Where variants differed was in schematic organization and coding as text or image. The PP scheme is a serial system, with a proper entry point and one prescribed reading order. In contrast, FI is a parallel system with any number of entry points and no prescribed order for addressing its elements. A result suggesting comprehension superiority for PP might suggest that FI is too much for 7th grade students to handle. But they do appear to have the faculties to navigate these complex visual environments and extract information from them, including conceptual information. Visual literacy is important in science (Akaygun & Jones, 2014; Ainsworth, 1999; Lowe, 2000). The comprehension test results suggest that as of the 7th grade, students have a demonstrable level of visual literacy, as they are able to succeed with high levels of text–image integration.

Media interest battery

Media interest data collection

The media interest battery utilized the online survey tool SurveyGizmo. Each media variant spread was broken into two separate images: a left-hand page and a right-hand page. Subjects made separate selections of "most interesting" for positive items, and "least interesting" for negative items, on different screens, by clicking on one of the images, which then displayed a check mark. If an opposing image were subsequently clicked, the check would move to the more recently selected image. Subjects had to complete all positive items before progressing onto the negative items screen, where they were given different prompts for the same sets of images. Page images were shown as thumbnails to ensure that subjects were reacting to the gist of the media variants (Carroll et al., 1992). This simulates a first impression of media and thus gauges situational interest. The displayed order of both the items within each screen and the options for each item (PP, PS, FI) were randomized per subject by the software.

Table 1. Comprehension test score means per media variant

Session	Media Variant	n	Mean
Session #1	Prose Primary (PP)	47	5.8085
	Prose Subsumed (PS)	51	5.9608
	Fully Integrated (FI)	48	5.9792
	All media variants	146	5.9178
Session #2	Prose Primary (PP)	54	7.0926
	Prose Subsumed (PS)	43	8.1395
	Fully Integrated (FI)	46	8.5435
	All media variants	143	7.8741
Session #3	Prose Primary (PP)	44	7.3636
	Prose Subsumed (PS)	53	6.9245
	Fully Integrated (FI)	51	7.4902
	All media variants	148	7.2500

Due to a glitch related to image file names, the left-hand pages used in the first comprehension test failed to display for both positive and negative items. Thus there were ultimately 5 positive and 5 negative items rather than the planned 6 and 6.

Media interest data analysis and reliability

In order to gauge the internal reliability of the media interest selections, a preference score was generated from item totals for each subject. For positive items, a selection of FI was coded as a point (+1) towards text–image integration, PS as zero (+0), and PP as a point deduction (−1). Negative items were reverse coded. Cronbach's alpha was calculated using the preference scores. The media interest data was rated reliable ($\alpha=0.8649$, $N=137$).

Unadjusted interest scores per text–image integration scheme were used to gauge the significance of the results. Individual interest scores were generated for each scheme. Any selection of a text–image integration scheme as most interesting increased its strategy interest score (SI) by one point (+1) from a base value of 5. Any selection of the same strategy as least interesting decreased its SI score by one (−1). When other strategies were selected as most or least interesting, the SI did not change (+0). Thus each subject had three scheme interest scores (in the range of 0…10), one for each of PP, PS, and FI.

True means, as 95% confidence intervals, were calculated to determine whether any differences in SI were significant.

Media interest results

Media interest selection results are reported in **Hata! Başvuru kaynağı bulunamadı.**. Across the positive interest items, selections favored fully integrated (FI: 64% of selections; PS: 20%; PP: 15%). Across the negative interest items, selections "favored" prose primary (PP: 52%; PS: 29%; FI: 20%). Thus the results suggest a general interest pattern of FI>PS>PP, which is consistent with the assumption that PS represents a mid-point between PP and FI.

True means estimates for strategy interest (SI) at 95% confidence are reported in **Hata! Başvuru kaynağı bulunamadı.**. Estimates are 3.1898 for SI-PP, 4.5693 for SI-PS, and 7.2409 for SI-FI ($N=137$). The confidence intervals reveal significant separation between all true means according to the interest pattern of FI>PS>PP. The subjects' interest increased in turn with increases in text–image integration, suggesting that the integration of text and image corresponds with greater situational interest.

There are unequivocal results for research question #2: How do different schemes for text–image integration in a science textbook impact middle school students' interest level in curriculum materials? Given the presentation of media variants for the media interest battery, it is considered a measure of text–image integration scheme's effect on situational interest. Chiappetta, Fillman, and Sethna (1991) suggested that research in science education should consider related factors beyond comprehension. Situational interest, a motivational factor, is here considered an aspect of the reader's impression of media, where comprehension is an aspect of the reader's engagement with media. By giving subjects only thumbnail views to compare text–image integration schemes directly for interest, the first impression with media was simulated: assessing the gist of the scene (Carroll et al., 1992). The learner's first impression of a textbook page is a factor of their motivation that is manipulable by designers and publishers. (Personal interest, conversely to situational interest, is developed over time [Chen, Darst, & Pangrazi,

2001].) The findings here strongly suggest that higher levels of text–image integration increase middle school learners' interest in curriculum materials.

Table 2. Interest in text–image integration scheme, item selection means (*N*=137).

Media Interest Item	Prose Primary (PP) Selections		Prose Subsumed (PS) Selections		Fully Integrated (FI) Selections	
	n	Proportion	n	Proportion	n	Proportion
Positive Interest Item #1.2	20	0.1460	28	0.2044	89	0.6496
Positive #2.1	25	0.1825	45	0.3285	67	0.4891
Positive #2.2	9	0.0657	26	0.1898	102	0.7445
Positive #3.1	39	0.2847	17	0.1241	81	0.5912
Positive #3.2	12	0.0876	23	0.1679	102	0.7445
Positive Interest Mean	21.0	0.15	27.8	0.20	88.2	0.64
Negative Interest Item #2.1	76	0.5547	32	0.2336	29	0.2117
Negative #2.1	70	0.5110	37	0.2701	30	0.2190
Negative #2.2	93	0.6788	22	0.1606	22	0.1606
Negative #3.1	9	0.0657	100	0.7300	28	0.2044
Negative #3.2	105	0.7664	7	0.0511	25	0.1825
Negative Interest Mean	70.6	0.52	39.6	0.29	26.8	0.20

Table 3. True means estimates for scheme interest scores (*N*=137).

Strategy	Parameter	Estimate	Confidence Intervals (95%)	
			Lower	Upper
Prose Primary (PP)	Mean	3.1898	2.8352	3.5444
	Std. Dev.	2.0988	1.8762	2.3817
Prose Subsumed (PS)	Mean	4.5693	4.2705	4.8682
	Std. Dev.	1.7690	1.5814	2.0075
Fully Integrated (FI)	Mean	7.2409	6.7236	7.7582
	Std. Dev.	3.0619	2.7373	3.4747

Given the research design, it is not possible to determine here if subjects were responding positively to more images, responding negatively to large passages of text, or both. However, though it is possible to separate such factors experimentally, in practice limited space is available for whatever content developers determine to be crucial information, and the provision of more images or diagrams does necessarily mean decreasing prose passages.

DISCUSSION

Limitations

This study was conducted with all 7th grade students at a middle school. It cannot be assumed that the results would hold for different age groups. It may be that as general literacy increases, prose primary (PP) becomes more appropriate. The ideal text–image integration scheme is likely not only a factor of the reader, but also content. As science material becomes more sophisticated, or more abstract, it may be more appropriately expressed through predominantly verbal means, as in PP media. Likewise, the branch of science should greatly impact the ideal integration scheme. Akaygun and Jones (2014) note how topics in science pose different challenges of visualization.

The media variants prepared for this study, as single isolated spreads from a science textbook, were relatively short on content. This limited exposure time to the materials and restricted the number of test items. As such the study was blind to any effects that might emerge with prolonged exposure. In particular, if fully integrated (FI) is a relatively novel format for the science textbook, then reading strategy, and ultimately reading comprehension, should improve over time through familiarity. Though this limitation has important theoretical implications, limited access at the participating school made extended interaction with media impossible. It is reasonable to assume that, in so much as FI is unconventional, increased practice in the scheme should increase its benefits. However, while FI may be unconventional as a general strategy for entire sections of science textbooks, diagrams are inherently FI, and students may have already developed strategies for reading in this scheme through repeated exposure to diagrams.

It is important to note that the media variants for this study were modifications from an original PP example. Ideally, instructional media content would be developed according to the text–image integration scheme to be employed. If the media for this study were developed from scratch with content experts, experienced science writers, and a designer, the resulting prose subsumed (PS) and FI media may have been more effective yet. However, the method employed did serve to make media variant equivalency easier to determine.

Implications

Image and schematic are influential in science education. Their use in the science textbook is especially important because textbooks play a definitional role in science education, as "delivered curricula" (Chiappetta & Koballa, 2002). A survey of American textbooks two decades ago found that images tended to be decorative or otherwise offer no new information (Woodward, 1993b). A more recent survey found that iconic photography is replacing explanatory or schematic images, an unfortunate trend (Lee, 2010). By favoring authentic media over simplified stimuli, this study went beyond a comparison of text and image to address information in the FI scheme in diagrammatic form.

Researchers in science education have explored the importance of the image in curriculum materials in specific subject areas since the use of images in textbooks was known to be problematic (Akaygun & Jones, 2013; Mayer et al., 1995; Peña & Quílez, 2001; Stylianidou, 2010), but only occasionally is text–image integration directly measured, especially in general schematic formats. Mayer et al. (1995) made an especially useful contribution by measuring the effects of prose-to-illustration proximity as well as the textual annotation of images. The differences between proximity and annotation, however, are only meaningful in PP and related schemes. The present study used a holistically schematic FI scheme that suggests new component descriptions (rather than Mayer et al.'s proximity and annotation). The results suggest benefits to this scheme.

The North American textbook production model (DiGiuseppe, 2014) poses an obstruction to increasing text–image integration in science textbooks, as content is developed in textual format before designers are involved in a subsequent production stage. And as evident from the production of the media variants for this study, FI media take much longer to design than their traditional PP counterparts. Together, the North American textbook production model and the design demands of FI media increase the burden of proof to change design practices: the benefits must outweigh the costs to prove practical. But if text–image integration is critical in science curriculum materials, efforts to that end are worthwhile.

In determining when and where a higher-integration scheme should be employed, subject areas (or scientific disciplines) will have different needs and

present unique challenges. Visualizing concepts in cell theory is not the same problem as visualizing concepts in Earth science. Akaygun and Jones (2014, citing Cheng & Gilbert, 2009) note that the ideal fit of information to text or image codes for specific science concepts is not currently understood. Fully integrated media, where code designation is compounded with schematic organization, complicates the issue further. Again, the results here suggest that exploring the fit of visualization strategy to science subject areas is a worthwhile pursuit.

Given the influence textbooks have in the classroom, however undesirable, it is important to understand how teachers use them. Would a shift towards higher text–image integration in the science textbook place different demands on teachers for which they might not be aware? In any case, considering the use of science curriculum materials and the involved and surely difficult-to-change textbook production model, the textbook as a format might be reconsidered. It is, after all, a series of lessons. Must these lessons be collected in an expensive bound volume? It may seem impractical to produce an entire textbook in an FI scheme. But content developers and researchers could make more compact materials available to teachers and school systems. One of the participating teachers in this study noted how desperate she was for activities and instructional media, saying that if she returns from a conference with just one new thing she can use in the classroom, she considers it a success.

This study provides a coherent framework for text–image integration for which the implications are likely broad. It is hoped that it can provide assistance in both the preparation and evaluation of curriculum materials in science education.

NOTES

[1] As an example of the conversion process, the "lakes & lake turnover" PP variant (**Hata! Başvuru kaynağı bulunamadı.**) included cross-section diagrams of a lake in both summer and fall with a paragraph describing winter. For the FI variant (**Hata! Başvuru kaynağı bulunamadı.**), the winter paragraph was converted into a winter diagram, which resulted in a decrease of total words from 74 to 31. The final phrase of the winter paragraph is: "The surface ice is near freezing, so the fish move to the bottom, where the water is a bit warmer." In the corresponding FI diagram, the equivalent text comes in three chunks arranged in two loci: (1) "fish prefer the warmer water," near the bottom of the lake and with a picture of fish, and (2) "water near freezing" placed just below the level of the ice, with a dotted line below separating it from the lower "water a bit warmer." The PP variant's proposition, that the bottom of the lake is "where the water is a bit warmer," becomes dual or split-coded in the FI variant, where water being warmer appears in textual form ("water a bit warmer"), but not *bottom of the lake.* Connecting warm water to the bottom of the lake is visual. The textual chunk does appear slightly low in the representation of the lake cross section. But the reader must recognize this positioning as meaningful. The dotted line serves this purpose, establishing an above/below relationship between "water near freezing" and "water a bit warmer."

[2] Given the low item count, individual factual and conceptual sets themselves (when not combined) had lower reliability estimates and did not achieve significance.

ACKNOWLEDGMENTS

REFERENCES

Akaygun, S., & Jones, L. L. (2014). Words or pictures: A comparison of written and pictorial explanations of physical and chemical equilibria. *International Journal of Science Education 36*(5), 783–807. doi:10.1080/09500693.2013.828361

Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education 33*, 131–152.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology 87*, 559–575.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives (complete edition).* New York, NY: Longman.

Asher, S. R. (1980). Topic interest and children's reading comprehension. In Spiro, R., Bruce, B. C., & Brewer, W. F. (Eds.), *Theoretical Issues in Reading Comprehension,* Erlbaum, Hillsdale, NJ, pp. 525–534.

Baddeley, A. (1998). Working memory. *Comptes Rendus de l'Académie des Sciences, 321*(2–3), 167–173.

Carroll, P. J., Young, J. R., & Guertin, M. S. (1992). A visual analysis of cartoons: A view from "The Far Side." In K. Rayner, (Ed.), *Eye movements and visual cognition* 444–461. New York, NY: Springer-Verlag.

Chambliss, M. J., & Calfee, R. C. (1998). *Textbooks for learning: Nurturing children's minds.* Maiden, MA: Blackwell.

Chen, A., Darst, P. W., Pangrazi, R. P. (2001). An examination of situational interest and its sources. *British Journal of Educational Psychology 71*, 383–400. doi:10.1348/000709901158578

Chiappetta, E. L., Fillman, D. A., & Sethna, G. H. (1991). A method to quantify major themes of scientific literacy in science textbooks. *Journal of Research in Science Teaching 28*(8), 713–725.

Chiappetta, E. L., & Koballa, T. R. (2002). *Science instruction in the middle and secondary schools.* Upper Saddle River, NJ: Merrill/Prentice Hall.

Dall'Alba, G., Walsh, E., Bowden, J., Martin, E., Masters, G., Ramsden, P., & Stephanou, A. (1993). Textbook treatments and students' understanding of acceleration. *Journal of Research in Science Teaching 30*(7), 621–635. doi:10.1007/BF02356924

Dechsri, P., Jones, L. L., & Heikkinen, H. W. (1997). Effect of a laboratory manual design incorporating visual information-processing aids on student learning and attitudes. *Journal of Research in Science Teaching 34*(9), 891–904.

DiGisi, L. L., & Willett, J. B. (1995). What high school biology teachers say about their textbook use: A descriptive study. *Journal of Research in Science Teaching 32*(2), 123–142.

DiGiuseppe, M. (2014). Representing nature of science in a science textbook: Exploring author-editor-publisher interactions. *International Journal of Science Education 36*(7), 1061–1082. doi:10.1080/09500693.2013.840405

Finson, K., & Pederson, J. (2011). What are visual data and what utility do they have in science education? *Journal of Visual Literacy 30*(2), 20–39.

Good, R. (1993). Editorial: Science textbook analysis. *Journal of Research in Science Teaching 30*(7), 619.

Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review, 1*, 69–82. doi:10.1016/j.edurev.2006.09.001

Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning and development. In Renninger, A., Hidi, S., & Krapp, A. (Eds.), *The role of interest in learning and development* (pp. 3–25). Hillsdale, NJ: Erlbaum.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education 33*(1), 27–50. doi:10.1080/09500693.2011.518645

Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching 44*(5), 706–724. doi:10.1002/tea.20172

Lee, V. R. (2010). Adaptations and continuities in the use and design of visual representations in US middle school science textbooks. *International Journal of Science Education, 32*(8), 1099–1126. doi:10.1080/09500690903253916

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology Journal, 30*(4), 195–232. doi:10.1007/BF02765184

Levin, J. R. (1979). On functions of pictures in prose. In *Report from the project on studies in language: Reading and communication*. Madison, WI: Wisconsin Research and Development Center for Individualized Schooling.

Lin, H., Hong, Z., & Chen, Y. (2013). Exploring the development of college students' situational interest in learning science. *International Journal of Science Education, 35*(13), 2152–2173. doi:10.1080/09500693.2013.818261

Lowe, R. K. (2000). Visual literacy and learning in science. *ERIC Digest, 3–4.* SE064310 EDO-SE-00-02.

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text?: Evidence from eye-movement patterns. *Computers and Education, 60*, 95–109. doi:10.1016/j.compedu.2012.07.011

Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Mayer, R. E., and Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology 82*(4): 715–726.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology and Research Development 43*(1), 31–43. doi:10.1007/BF02300480

Miller, J. S., & Krumhansl, R. (2009). Learning from innovative instructional materials and making them your own. In Gess-Newsome, J., Luft, J. A., & Bell, R. (Eds.), *Reforming secondary science instruction* (pp. 39–52). National Science Teachers Association. Arlington, VA: NSTA Press.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology 85*, 424–436.

National Education Goals Panel (1998). *Recommendations regarding the implementation of standards.* Retrieved June 13, 2015, from govinfo.library.unt.edu/negp/page1-13-9.htm.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance and monitoring accuracy. *Metacognition and Learning, 2*, 159–179. doi:10.1007/s10409-006-9595-6

Peña, B. M., & Quílez, J. M. G. (2001). The importance of images in astronomy education. *International Journal of Science Education 23*(11), 1125–1135. doi:10.1080/09500690110038611

Pettersson, R. (1998). Image functions in information design. The 30th annual conference of the International Visual Literacy Association, University of Georgia, Athens, Georgia.

Plass, J. L., Moreno, R., & Brünken, R. (Eds.) (2010). *Cognitive load theory.* Cambridge: Cambridge University Press.

Pozzer, L. L., & Roth, W. M. (2003). Prevalence, function, and structure of photographs in high school biology textbooks. *Journal of Research in Science Teaching 40*(10), 1089–1114. doi:10.1002/tea.10122

Reynolds, C. R., Livingston, R. B., & Willson, V. (2006). *Measurement and assessment in education.* Boston: Pearson.

Roseman, J. E., Stern, L., & Koppal, M. (2010). A method for analyzing the coherence of high school biology textbooks. *Journal of Research in Science Teaching 47*(1), 47–70. doi:10.1002/tea.20305

Rotgans, J. I., & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction 32*, 37–50. doi:10.1016/j.learninstruc.2014.01.002

Roth, W. M., Bowen, G. M., & McGinn, M. K. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching 36*(9), 977–1019.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Schiefele, U. (1991). Interest, learning and motivation. *Educational Psychologist, 26*, 299–323. doi:10.1080/00461520.1991.9653136

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19*, 469–508. doi:10.1007/s10648-007-9053-4

Schraw, G. (1997). Situational interest in literary text. *Contemporary Educational Psychology, 22*, 436–459.

Schraw, G., & Lehman, S. (2011). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review 13*(1), 23–52. doi:10.1023/A:1009004801455

Stern, L., & Roseman, J. E. (2004). Can middle-school science textbooks help students learn important ideas? Findings from Project 2061's curriculum evaluation study: Life science. *Journal of Research in Science Teaching 41*(6), 538–568. doi:10.1002/tea.20019

Swarat, S., Ortony, A., & Revelle, W. (2012). Activity matters: Understanding student interest in school science. *Journal of Research in Science Teaching 49*(4), 515–537. doi:10.1002/tea.21010

Tapola, A., Veermans, M., & Niemivirta, M. (2013). Predictors and outcomes of situational interest during a science learning task. *Instructional Science 41*, 1047–1064. doi:10.1007/s11251-013-9273-6

Trefil, J, Calvo, R. A., & Cutler, K. (2005). McDougal Littell science, North Carolina edition, student edition, course 3. Evanston, IL: Houghton Mifflin Company.

Trend, R. (2005). Individual, situational and topic interest in geoscience among 11- and 12-year-old children. *Research Papers in Education 20*(3), 271–302,

Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science, 3*, 499–535. doi:10.1111/j.1756-8765.2010.01113.x

U. S. Department of Education [USDOE] (2010). *Public elementary/secondary school universe survey, 2000–01, 2005–06, 2007–08, and 2008–09.* National Center for Education Statistics, Common Core of Data (CCD).

Van Genuchten, E., Scheiter, K., & Schüler, A. (2012). Examining learning from text and pictures for different task types: Does the multimedia effect differ for conceptual, causal, and procedural tasks? *Computers in Human Behavior, 28*, 2209–2218. doi:10.1016/j.chb.2012.06.028

Vekiri, I. (2002). What is the value of graphical displays in learning?. *Educational Psychology Review, 14*(3): 261–312. doi:10.1023/A:1016064429161

Wake County Public School System [WCPSS] (2011). *Demographics 2010–2011.* Retrieved August 15, 2014, from www.wcpss.net/about-us/our-students/demographics/reports/book10.pdf.

Woodward, A. (1993a). Introduction: Learning from textbooks. In B. K. Britton, A. Woodward, & M. Binkley (Eds.), *Learning from textbooks: Theory and practice* (pp. vii–x). Hillsdale: Lawrence Erlbaum Associates Publisher.

Woodward, A. (1993b). Do illustrations serve an instructional purpose in U. S. textbooks? In B. K. Britton, A. Woodward, & M. Binkley (Eds.), *Learning from textbooks: Theory and practice* (pp. 115–134). Hillsdale: Lawrence Erlbaum Associates Publisher.

Worthy, J. (1996). A matter of interest: Literature that hooks reluctant readers and keeps them reading. *The Reading Teacher 50*, 204–212.

❖ ❖ ❖