International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Alternate text types and student outcomes: an experiment comparing traditional textbooks and more epistemologically considerate texts

Matthew Kloser

To cite this article: Matthew Kloser (2016): Alternate text types and student outcomes: an experiment comparing traditional textbooks and more epistemologically considerate texts, International Journal of Science Education, DOI: 10.1080/09500693.2016.1249532

To link to this article: http://dx.doi.org/10.1080/09500693.2016.1249532

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
http://dx.doi.org/10.1080/09500693.2016.1249532

Routledge
Taylor & Francis Group

Alternate text types and student outcomes: an experiment comparing traditional textbooks and more epistemologically considerate texts

Matthew Kloser

Institute for Educational Initiatives, University of Notre Dame, Notre Dame, IN, USA

ABSTRACT

Inscriptions and texts are central to the practice of science and determining how science ideas are presented in high school biology classrooms. Traditional textbooks have been criticised for their expository nature, their difficult lexical structure, and for their lack of evidence to support claims. Recent frameworks for science education in the United States emphasise the need for students to engage texts that better reflect the scientific enterprise. This study uses a randomised experiment to compare high school student outcomes – interest, comprehension, and learning – when reading traditional biology text accounts and when reading more epistemologically considerate (EC) accounts that provide developmentally appropriate narratives of scientific experiments, including data and evidence for claims made within the texts. Results indicate that students in the two conditions may not differ in their interest in or comprehension of the texts, but that students reading the more EC texts show some higher achievement on transfer tasks. Furthermore, students in the treatment condition show that when epistemic resources are made available within a text, they are more likely to attend to data and evidence and are less likely to rely on the authority of the text when determining the trustworthiness of claims.

ARTICLE HISTORY
Received 12 May 2016
Accepted 13 October 2016

KEYWORDS
Textbooks; epistemology; biology education

Young peoples' images of science often portray the stereotype of old men working with chemicals in an isolated laboratory setting. These images focus on experimentation and lab work as the central function of science (Finson, 2002). In contrast, Latour's (1986) anthropological exploration of a science lab portrays the enterprise as the constant interplay among groups of people and various forms of text. Inscriptions, graphs, and journal articles are central to the theoretical and empirical development of science ideas. In American science classrooms, texts – particularly the traditional textbook (TT) – have a similar privilege (Weiss, Pasley, Smith, Banilower, & Heck, 2003). If children were asked to draw images of school science, their pictures may very well include textbooks as central to what occurs in the classroom.

CONTACT Matthew Kloser mkloser@nd.edu Institute for Educational Initiatives, University of Notre Dame, 107 Carole Sandner Hall, Notre Dame, IN 46556, USA

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/09500693.2016.1249532

Unfortunately, Latour's inscription-based portrayal of bench science and an image of textbook-based school science are not equivalent in their representation of the discipline. Scientists engage texts that contribute to the epistemic function of the discipline; they read tables and visual representations of data used in the development of scientific arguments. Their methods are made explicit through the communication of peer-reviewed literature. Meanwhile, students exposed to TTs are presented with a 'rhetoric of conclusions' (Schwab, 1962, p. 24) that masks the evidentiary basis for the text's claims. While presenting the empirical support and justification for every scientific claim learned in a science classroom would be both inefficient and developmentally inappropriate, TTs have so thoroughly ignored the structure and the nature of the discipline that the extensive use of these resources can undermine what students learn *about* science.

Recent standards documents in the United States emphasise the use of expanded text types in science classrooms beyond the TT. The scientific practice of 'obtaining, evaluating, and communicating information' in the *Framework for K-12 science education* (National Research Council, 2012) connotes science literacy with 'the ability to read and understand [science's] literatures' (p. 74). This practice includes interpreting words, symbols, diagrams, graphs, and images, thus requiring more than the extraction of technical vocabulary, but rather precise meaning of ideas through multiple modes. The *Common core standards for English language arts & literacy* (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) are more specific in detailing 10 student outcomes for students' engagement with texts in science and the technical subjects. Among them, students should be able to trace a text's depiction of a complex process or phenomenon (standard 10.2). They should also be able to describe experimental procedures and identify a scientist's research questions (standard 10.6) while assessing whether the reasoning and evidence support an author's claims (standard 10.8).

Texts play a major role in science and science classrooms, but more research is needed to understand how alternative text types may affect student understanding and interest both of and about science ideas. This study uses custom-designed texts in a randomised trial with high school biology students to compare student outcomes between TT accounts and accounts that make their claims and supporting evidence more epistemologically considerate (EC). The EC accounts reflect more of the structure of the discipline by using questions as headings, rather than expository statements, by embedding developmentally appropriate representations of authentic data alongside scientific claims, and by placing the science ideas within their historical context. The following section describes the current use and structure of textbooks that dominate biology classrooms followed by a conceptual model for understanding how students learn and engage texts more broadly, placing the results of this study in context.

Background literature

TTs are incredibly efficient in representing decades and even centuries of scientific understanding in a single volume. Science builds on existing theories and knowledge and progresses because previous conclusions are available to future generations. This efficiency, however, results in trade-offs. As Schwab (1978) commented nearly 40 years ago, 'It is a rare textbook, indeed, which supplies enough of the structure of the discipline to let students know that he is dealing with a model or a possibility and not with a literal truth or literal falsehood' (p. 236). TTs generally provide type 5 claims (Latour, 1986) – certitudes that are taken authoritatively and do not make explicit the messy data or productive struggle of the research that helps inform these conclusions. In the United States, the textbook has been described as the single most influential tool used by the teacher to determine a class's curriculum and instruction (Hurd, Bybee, Kahle, & Yager, 1980), thus passing on to students a wide range of unjustified claims.

Science textbook use varies among classrooms, but three decades of literature indicate that textbooks, especially for novice teachers, are key planning and instructional tools (Tobin & Gallagher, 1987; Tulip & Cook, 1993; Weiss et al., 2003). In a national survey, biology teachers self-reported that about 50% of what is taught and over 70% of how it is taught is based on the textbook (Weiss et al., 2003). Tulip and Cook's (1993) examination of science textbook usage suggested that the textbook decided for many teachers what, when, and to what depth science content should be taught.

Science textbook content and style

TTs have the potential to guide how students view a discipline and its interaction with the world (Chambliss & Calfee, 1998). The structure and content of biology textbooks, and science textbooks more broadly, thus carries significance. On average, science textbooks in the United States contain more pages than textbooks from other disciplines (Chiappetta & Fillman, 2007). These volumes contain extensive new and technical vocabulary that in biology texts surpasses even the other sciences (Merzyn, 1987). In addition to an emphasis on vocabulary, biology textbooks generally share an expository writing style (Chiappetta & Fillman, 2007; Goldman & Bisanz, 2002; Phillips & Norris, 2009). Generally, these expository prose present scientific ideas as 'final form science' (Duschl, 1990) – authorless accounts, detached from context, that present students with unchallengeable facts (American Association for the Advancement of Science, 2005). The expository style, while efficient, allows students few opportunities to engage in the structure of the discipline, leaving many students to memorise abstract pieces of knowledge rather than place it within a larger conceptual framework (Myers, 1992).

TTs also contain visual representations that affect comprehension and other cognitive processes. In a meta-analysis comparing 75 texts with and without illustrations, Levin and Mayer (1993) found that texts with illustrations have a large positive effect ($d = 0.80$) on comprehension. They posited that visual representations provide concrete examples of the phenomena in question and help the reader improve coherence among ideas. While visual representations such as schematics, graphs, and charts serve an explanatory role, approximately 85% of images in science textbooks have no clear relevance to aiding comprehension and, at times, can increase student misconceptions (Lee, 2010). In sum, both the structure and content of traditional science textbooks, especially biology texts, create obstacles for student comprehension and learning of concepts while presenting a view of science that lacks evidence for stated claims.

Alternative science texts

TTs are not the only means by which students learn biology. Even in the 1960s when TTs were seen as the singular curricular resource, science educators proposed alternative text types that conveyed scientific accounts drawing on a research problem and accompanied by data (Olson, 2000). Historical accounts of science and adapted primary literature (APL) are two examples of alternative text types.

Historical accounts of science can provide insight into the socio-cultural events that are responsible for scientific claims and can help address students' naive conceptions of science (Brush, 1974; Matthews, 1994; Monk & Osborne, 1997). Historical accounts necessarily avoid type 5 claims and expose the practices and ideas of science that are eventually represented in the peer-reviewed literature and textbooks. By exposing the social processes of science, historical texts can introduce students to a variety of methodologies and model argument development over time. Most importantly, as Monk and Osborne (1997) stress, historical accounts can reveal elements of the epistemology of science, leading students to understand how we know certain ideas that we take as fact.

Research on the use of historical accounts of science in classrooms is limited, but another alternative text type that seeks to address elements of the discipline's epistemology, APL, has an emerging basis in the literature. While the major contributions of science are models, theories, and knowledge about the natural and material world, the currency of scientific practice is the peer-reviewed journal article. APL takes up the peer-reviewed journal article as an authentic text for student engagement by modifying elements in a developmentally appropriate way. APL adapts the syntax, sentence structure, and data representations to convey both content and process to high school and undergraduate students (Norris et al., 2009; Phillips & Norris, 2009; Yarden, Brill, & Falk, 2001). This adapted text exposes students to authentic research questions, a variety of methods, and the language and structure of scientific communication.

Studies indicate mixed results for the use of APL in classrooms. One study showed that high school students who read APL were more likely to critically discuss material and ask questions of higher cognitive levels compared to students who read conventional texts (Brill & Yarden, 2003). Another study showed that APL helped students learn aspects of the language of science, the epistemology of science, and how to identify misconceptions (Yarden et al., 2001). Neither study measured students' comprehension or ability to use the texts' ideas in new contexts. Criticisms of APL cite that many students lack the necessary background knowledge to comprehend and make connections within the text (Osborne, 2009). Furthermore, as a product of a single modified journal article, the use of APL could become an inefficient means for helping students learn a body of scientific knowledge within a given year (Norris et al., 2009).

TTs dominate the market and use in schools (Wood, 2002). Yet their structure and content have been criticised over multiple decades for forcing students into memorising facts and accepting claims without epistemic justification. While alternative texts have been proposed, some sceptics believe that these texts will focus too narrowly on individual ideas or reflect the peer-reviewed journal article structure that also will hinder student comprehension. The study reported here draws on the following conceptual framework to investigate whether alternative text types designed in a more narrative fashion, but that illuminate aspects of the structure of the discipline, can result in student comprehension and learning both of and about the discipline.

Conceptual framework

Text processing

Reading a text is a dynamic process that involves an interaction between attributes of the reader, attributes of the text, cognitive processes, and context. In science classrooms, the goals of this dynamic process are generally focused on comprehension and learning. Within the framework described below, comprehension represents instances in which a reader activates a schema that contains an appropriate account of the message contained in the text (Anderson, 1984). Learning includes comprehending the ideas of texts, but goes beyond understanding words, sentences, and paragraphs and their meaning (van den Broek, 2010). Learning requires that ideas from a text are integrated with existing knowledge in ways that the ideas can be used meaningfully (Otero, León, & Graesser, 2002). Thus, when a reader comprehends a text and can integrate this message with prior knowledge or shape existing understanding, learning can occur (Chambliss, 2002).

The act of comprehending and learning from texts requires cognitive engagement that processes the text. While reading a text, a reader has finite cognitive resources to devote to the task – resources that are impinged upon by other stimuli that surround the reader. Current theories of comprehension and learning from texts recognise the importance of attention allocation (van den Broek, 2010; van den Broek, Young, Tzeng, & Linderholm, 1999). As one reads, information varies in importance and interest and the focus of attention changes as connections are made between the text and one's prior knowledge. The elements of the text that receive the greatest attention have the highest probability of being understood and incorporated into one's prior knowledge. In this study, the EC texts use a more narrative structure detailing the people and processes involved in investigating a scientific idea. Compared to the expository structure of the TT account, the narrative form has been linked to greater reader interest and, thus, attention allocation (Ainley, Hidi, & Berndorff, 2002).

Attending to a text is necessary, but not sufficient to fully understand its meaning or use in new contexts. The reader must be able to semantically analyse its parts as well as generate a mental representation of its contents. Semantic analysis involves the decoding of words, sentences, and paragraphs. In science texts, this requires a unique ability to interpret extensive amounts of new vocabulary and navigate lexico-grammatical structures foreign to more common texts. Science writing is known for its passive writing style that lacks logical connectives between parts of the texts as well as its use of nominalisations and lexical metaphors (Norris & Phillips, 2003; Wellington & Osborne, 2001). If readers are able to navigate the semantic structure and vocabulary, processing a text is generally accompanied by the generation of a mental representation of the text (Chambliss & Calfee, 1998; McNamara, Kintsch, Songer, & Kintsch, 1996). A mental representation is an image in one's head that is modified as new information is presented, inferences are generated, and gaps in the mental model are filled (Maury, Perez, & León, 2002). High-quality mental representations have been shown to correlate with better comprehension of a text's meaning and better recall of the meaning in the future (van den Broek, 2010; van den Broek et al., 1999).

As mentioned above, TTs generally present ideas in an expository format that include a vast amount of new vocabulary. In general, expository science texts place a high cognitive burden on students' ability to form coherent, mental representations (Calfee & Curley, 1984; Elshout-Mohr & van Daalen-Kapteijns, 2002; Graesser, Singer, & Trabasso, 1994). While the EC and TT accounts in this study were controlled for reading level based on word and sentence length, the essential elements of the EC text – the more narrative structure and the graphical representations of data – were developed to better aid semantic analysis and the development of a mental representation. Given the narrative structure's contextualisation and logical sequencing, it is posited that readers will be better able to comprehend the words and phrases of the text (León & Penalba, 2002). Similarly, unlike the iconic images displayed in most TTs, the presence of data and figures that support the main claims of the text will theoretically allow a more robust mental representation of the text's ideas as the graphs and figures serve a stronger explanatory role.

Readers bring many resources to the reading experience as well. Prior knowledge about a text's ideas is an integral part of the process since 'comprehension requires background knowledge to construct a meaning of the text' and in learning, 'background knowledge becomes the object of change' (van den Broek, 2010, p. 453). Individuals with more background knowledge tend to create stronger mental representations that in turn facilitate recall and long-term retention. Furthermore, individuals not only have prior knowledge related to the content of a text, but also existing schemas of how to approach different text types. A schema aids readers in making predictions, interpreting, and evaluating texts. By the time students enter a high school biology classroom, they generally have developed a schema for reading traditional science textbooks accounts (Garner, 1987).

Although not comprehensive, one final important construct unique to each reader is the level of interest in a text type and its content. Interest levels are significant mediators of attention allocation and, thus, have an impact on the processing of a text. For example, reading a book on one's favourite historical period is more likely to be assimilated into one's prior knowledge and generate a helpful representation than the decontextualised reading of a refrigerator's user manual. Texts can trigger situational interest – that is, 'focused attention and [an] affective reaction to a stimulus in the environment that may or may not last over time' (Hidi & Renninger, 2006, p. 113). Studies have shown that novel ideas, narrative stories, and issues dealing with living things, among other things, tend to spark situational interest (Bergin, 1999). In a comparison of text types, interest levels, and comprehension, interest in a text has a small to medium effect on comprehension and narrative texts are found more interesting than expository texts (Schiefele, 1992). The role of interest in aiding comprehension and learning is only positive, however, if the stimulus inducing interest is aligned with the purpose of the text. Seductive details, such as bright coloured pictures or alarming details that do not help make connections between the message of the text and the reader's prior knowledge, can divert attention and reduce comprehension (Alexander & Jetton, 1996) – an important distinction for TTs in which the extensive visual representations often do not serve the purpose of connecting ideas within the text.

Scientific and personal epistemologies

The enterprise of science favours particular dispositions towards justifying claims with evidence. Scientists must make explicit their methods and ways of knowing so that other scientists and readers understand the epistemic foundations by which claims are made. Expertise may carry some caché – a Nobel Prize winning scientist is likely to convince many people just on their authority – but within the realm of publishing new ideas in journals, evidence and justification are required. As children engage in science, particular dispositions can be considered as more or less sophisticated epistemic stances. Accepting a claim based on an authoritative presence represents a less sophisticated scientific epistemology. The development of an argument supported by evidence represents a more sophisticated epistemology.

Recent research indicates that personal epistemologies are neither fixed nor always follow a particular developmental trajectory. A resource-allocation model of personal epistemologies suggests that students have a variety of epistemological resources ingrained from their experiences with the world, but need certain cues to activate more sophisticated ways of viewing knowledge and the construction of knowledge (Hammer & Elby, 2003; Louca, Elby, Hammer, & Kagey, 2004). In one context, students might reveal a less sophisticated epistemic stance – a naive acceptance of a claim on authority. But with the right cues and context, from a text, a teacher, or a classmate, the same child can exhibit a quite sophisticated stance that calls for substantial evidentiary and theoretical backing for claims. In dealing with texts specifically, a recent study gave students four different text types that contained different epistemological resources. When given resources of higher epistemic value, like justifications with evidence, students were more likely to choose these accounts as more reliable over accounts that simply made authoritative claims (Sandoval & Çam, 2011). Thus, the nature of the text and its components may not only affect comprehension and learning, but may also shape a reader's activation of a more or less sophisticated epistemological stance.

Research questions

The EC texts used in this study were designed to help students make connections within the text to their prior knowledge and use evidence to justify embedded claims. Furthermore, the texts were designed to better reflect aspects of the structure of biology as a discipline and provide students with empirical data that could help shape more scientifically sophisticated epistemic perspectives. Given the novelty of these texts, empirical testing that compares these accounts with TT accounts is necessary. To this end, I ask the following research questions:

(1) What is the effect of text type on student interest in reading science texts?
(2) What is the effect of text type – TT and EC – on student comprehension of science ideas?
(3) What is the effect of text type on student learning as evidenced by their ability to apply concepts to new biological situations?
(4) How do students use evidence from texts in justifying claims?

Methods

This study investigated how traditional biology textbook accounts (control condition) and more EC accounts (treatment condition) affect students' interest in, comprehension of, and learning of biology concepts. The EC and TT accounts were designed with structures and content that conceptually would lead to different attention allocation, semantic analysis, and the development of mental representations, thus affording different levels of comprehension, learning, and the activation of epistemic resources. Data were collected using a 2×2 (text condition by prior achievement in English language arts (ELA)) multivariate experimental design with a control group (Campbell & Stanley, 1966) and repeated measures across two vectors of outcomes, interest and cognition. This study also investigated the type of epistemological resources that readers in the two conditions activated when engaging the different text types.

Students in high school biology courses were stratified into two levels of prior achievement based on a median split of their eighth grade ELA scores, the common prior achievement metric in reading among all students. Prior achievement in ELA was used as a factor in order to see whether this text-based intervention had a differential impact on stronger and weaker readers. For example, it could be hypothesised that the presence of data from figures that support the texts' claims could aid students more who struggle with the semantic analysis of expository prose. Thus, by using prior ELA achievement as a factor, interactions between treatment condition and reading level could be explored.

Participants were randomly assigned to control and treatment conditions within each participating classroom. Students remained in their assigned condition across four replications of the experiment. An omnibus chi-squared test for gender, class year, ethnicity, English language learning (ELL) status, eighth grade ELA, and seventh grade science state standardised test scores indicated no significant differences between the treatment and control samples.

Following randomisation, students read their respective text type. Data were collected on their interest in the text, their comprehension of the text, and their ability to transfer information from the text to new situations. Outcomes were measured using three instruments: a question guide, an interest survey, and a content-focused post-test, all described below.

Setting and participants

Classrooms from two high schools in the United States with diverse populations were recruited to participate in the study, Rolling Hills High School (RH) and Bell View High School (BV).[1] RH and BV both serve a majority of students from underrepresented minority backgrounds. Forty-two per cent and 32%, respectively, of students from these schools receive free or reduced lunch. Students from three teachers' biology classrooms participated in the study. Although teachers played no role in the experiment, a survey of teacher practice indicated no difference in the rate at which the teachers discussed the nature of science, historical experiments, experimental design, or the use of the biology textbook in class. Selected classes ranged from college-prep biology to honors-level biology.

Two hundred and seventy students returned consent forms. From this pool of students, each class was randomly assigned to half the students in the control and half the students in the treatment conditions. Thirty-five of the students never participated in a single replication due to movement in their class schedules. Absences on the day of each replication resulted in minor fluctuations of student participation. The total sample included slightly more male than female students and Hispanic and White students represented over three-fourths of the participant sample (Table 1). Chi-squared tests for each replication indicated no difference in prior conditions for the treatment and control groups even after attrition.

Materials

Four text topics were chosen for this experiment based on the following criteria: (1) the topics included both microscopic and macroscopic concepts; (2) the concepts were present in widely used biology textbooks (such as Holt's *Modern Biology*); (3) the concepts were not explicitly covered by previous middle school life science standards; (4) the concepts could be introduced in approximately 500 words. Pairs of control and treatment texts were created based on TT accounts and historical experiments/journal articles, respectively. Six pairs of texts were piloted with local high school students. Four pairs of texts were chosen for this study and replications of the experiment are reported here for three texts; the fourth and final replication compromised unpredictable school events. The three pairs of texts focused on (1) cell signalling; (2) pancreatic physiology; and (3) social group ecology.

Control and treatment texts covered the same biological topic and major concepts in two distinct formats (see Table 2 for text component comparison). Control texts were written in a traditional expository, informative style. Headings were stated in an expository fashion and the main text used a conventional, fact-based style that supported concepts with examples and supporting details. Visual representations – pictures or drawn figures – were also embedded in the texts. In contrast, treatment texts incorporated historical accounts of seminal experiments. Section headings were written as research questions, rather than declarative statements. The treatment text contained a variety of statement types that helped build a narrative of the hypotheses and investigations conducted by scientists to address a research question. In each treatment text, evidence in the form of graphs or qualitative biological assays was used to support or negate the stated hypotheses. The visual representations of data were taken from the original experiments, but were simplified in their presentation to be developmentally appropriate (e.g. references to statistical significance and error bars were removed). Word count, readability as judged by the Flesch–Kincaid Grade Level formula, and the number of visual representations were controlled as much as possible between the two conditions. Word count never varied by more than 10% among the pairs, readability never differed by more than two-tenths of a point – the equivalence to two months of education – and the number of visual representations never varied by more than one image. A fourth-year doctoral candidate in immunology validated biology content for both text versions across each replication. Equivalence of the texts and assessments were validated by four professors – two professors of science education, one educational psychologist/psychometrician, and one professor of English education and critical theory. Examples of the two text types are available in Kloser (2013).

Reading response sheets contained room for students to record all questions that arose while reading the texts. Instructions required that all students write at least one

Table 1. Participating student demographics by school.

	Gender		Ethnicity					Class year			ELs	Eight grade ELA achievement			Seventh grade science achievement		
	M	F	Hispanic	White	Black	Asian	Other	Fr.	So.	Jr.		Below basic	Basic	Advanced	Below basic	Basic	Advanced
RHHS	95	75	103	39	15	9	6	22	108	38	40	28	56	72	48	61	39
BVHS	47	53	44	41	2	5	6	0	102	0	19	6	15	71	12	31	48
Totals	142	128	147	80	17	14	12	22	210	38	59	34	71	143	60	92	87

Table 2. Text element comparison between EC and TT accounts.

Text element	EC accounts	TT accounts
Headings	Questions	Expository statements
Prose components	Narrative and argumentative	Expository
Visual representations	Tables, charts, or graphs	Iconic photos or schematics
Topic	Same conceptual topic – see Appendix in supplementary file for detailed concept breakdown	
Length	Texts within 10% of each other's length	
Reading level	Reading levels within 0.2 of each other on the Fleish–Kincaide scale	

question about the text. The number of questions was used as a secondary measure of student interest – the more questions that students raised, the more engaged they were in either comprehending elements of the text or exploring ideas beyond the content of the texts. The backside of the reading response guide contained a five-question interest survey. Each survey question used a five-point Likert-style scale. A uni-polar scale ranging from 'Not Interested' to 'Extremely Interested' was used to measure student interest in the texts. Post-tests contained both multiple-choice and free-response questions. Four questions probed students' comprehension of terms and concepts addressed in both texts.

Two additional multiple-choice questions asked students to identify from a list of options a major claim found in the text. Each of these multiple-choice questions was followed by a free-response question asking students to identify evidence from the text that supported the given claim. For example, on the cell-signalling test, students were asked:

1 (a) Which of the following claims is a main idea from the text?
   - (a) Cells within the embryo send signals that cause cells to form different cell types
   - (b) Cells within the embryo differentiate based on how much sunlight they receive
   - (c) Cells within the embryo send signals that help specialised cells grow larger
   - (d) Cells within the embryo are all different and require light and heat to grow

1 (b) What evidence supports the claim you chose in Question 1(a)? (Please write your answer below).

Three final questions used a similar combination of multiple-choice and free-response items to measure students' abilities to transfer concepts from the texts to new situations. Students chose from a list of answers and had to justify their answer in a free-response section. Finally, the last question measured students' ability to design an experiment to test a claim related to the text's topic. For example, on the pancreatic physiology test, students were given the following instructions:

> A scientist has discovered a small molecule produced by the pancreas. He thinks this small molecule is a hormone that causes pigs to get thirsty. Please write an experiment you could run to test whether this molecule causes pigs to get thirsty.

Procedures

The experiment occurred over four weeks of instruction, prior to when the related topics would be taught in class. Students received an individualised packet with a unique ID number, respective of their assigned condition. Students read the assigned text individually. Upon completing the reading, students completed a reading response sheet. On the front side, students were asked to write down at least one question that arose while they read, but were encouraged to write down as many questions as came to mind. On the back, students completed the five-question interest survey. These materials were collected and all students received the same achievement post-test with their unique ID number. Each replication required an average of 30 minutes. A group of 24 students from the same student population, not included in the experiment and selected to balance gender, reading level, and ethnicity, participated in think-alouds with each of the texts. The results from these think-alouds are detailed in Kloser (2013) and referenced in the discussion below.

Data analysis

Interest scores were analysed as the dependent variable in a 2×2, treatment by prior achievement in ELA (median split) MANOVA to determine if text type resulted in different student interest levels. The outcome measures for the MANOVA included the interest survey scores and the number of questions that students asked on the reading guide. Both the survey and the number of questions sought to measure students' interest levels in the different texts. Levene's statistic for each replication indicated homogeneity of variance and a K–S test indicated that the interest survey data were normally distributed. For the interest survey questions, Cronbach's alpha ranged from 0.857 to 0.907 across the three texts. Given the high internal reliability, the values from the five questions were summed and treated as a composite interest measure.

Two independent raters scored achievement post-tests, blind to condition. Multiple-choice questions were scored according to a key. Free-response questions were coded using pre-determined codes. Agreement on free-response question coding exceeded 80% on a 20% sample for each of the three replications. The remaining tests in the sample were coded individually with periodic calibration checks. Free-response questions were of two types: (1) content questions that were included in the final post-test score and (2) 'evidence questions' that did not figure in the final score. Content questions were scored using a rubric of 0–2 points. Multiple-choice and content-focused free-response questions were analysed as dependent sub-scores – comprehension questions and transfer level questions – in a 2×2 (treatment condition by reading level) MANOVA. Cronbach's alpha for internal reliability across the content questions on the test ranged from a low of 0.743 to a high of 0.774 across the three replications of the experiment.

Evidence questions were coded by the raters using a taxonomy based on increasingly sophisticated epistemologies (Table 3) (Hammer & Elby, 2003; Louca et al., 2004). More naive epistemologies were represented by student responses that cited no evidence, relied on their own invented explanations, or cited the text's unjustified claims as an authoritative source. Increasingly sophisticated citations of evidence included

Table 3. Level 1 coding scheme for evidence questions asked by students.

Category	Code description	Example
Knowledge	The question asks about a fact or vocabulary term that can be answered directly from a common reference source such as a dictionary or science textbook glossary	What is a flock?
Comprehension	The question asks about concepts that can be answered by carefully reading the given text with minimal inference	Why do animals often live in a flock or pack?
Elaboration	The question asks about an extension of the text's concepts that require further information or significant inference	How do they test detection of the hawks by the pigeons?
Application	The question asks about how a theme or concept from the text can be applied to a new context or model system	Does the safety-in-numbers hypothesis apply to all other animals?

explanations or examples from the text. The code representing the most sophisticated response was applied when students cited empirical evidence to support a claim. Percentages of students' responses, since slightly more students in the treatment condition submitted post-tests, were compared using a split-plot ANOVA with the experimental condition as the between-subjects factor and the five levels of the taxonomy as the within-subjects repeated measure. In each case, Mauchly's test indicated that the assumption of sphericity had not been violated. *Post-hoc* analyses were performed for significant split-plot ANOVAs using a Bonferonni adjustment to determine which categories were significant between the two conditions.

Results

Interest

Results from the MANOVA focused on the interest measures – the interest survey score and the number of questions asked – indicated minimal statistically significant findings across the three replications. Interest measures by experimental condition showed no significance and only Text 1 showed an experimental condition × ELA prior achievement interaction, $F(2) = 4.26$, $p < .05$. Univariate results for the interaction effect for Text 1indicated that only the interest survey score was significant, $F(1) = 7.52$, $p < .01$. Estimated marginal means showed that for Text 1, students designated as high achievers after the median split for ELA prior achievement reported increased interest when they read the treatment text ($M = 3.02$, SE = 0.095) compared to when they read the control text ($M = 2.56$, SE = 1.09). The number of questions that students asked in the two conditions was not statistically different.

Comprehension and learning

MANOVA results across all three replications showed no significant differences in student performance on the comprehension level questions between the two experimental conditions ($p > .05$ for all three replications). Students in the top half of the ELA median split scored statistically higher on the comprehension level questions than students from the lower half of the median split. Unsurprisingly, this trend occurred across all three text replications ($p < .05$). However, no interactions occurred, suggesting that neither condition differentially helped students of different prior ELA achievement levels better comprehend the text.

Although the experiment resulted in no differences in comprehension scores, MANOVA results indicate significant differences for other cognitive measures for Text 1, $F(4) = 3.17$, $p < .05$ and Text 2, $F(4) = 4.00$, $p < .01$ (Figure 1). Univariate tests were run for Texts 1 and 2 to determine the specific cause of the differences. For Text 1, the experimental condition was significant for student performance on the experimental design transfer question block, $F(1) = 4.70$, $p < .05$. Estimated marginal means for the experimental design question indicate that across all three texts, students in the treatment condition scored higher than students in the control condition on these items, but only in Text 1 was this difference significant at the .05 level. For Texts 2 and 3 this difference was significant only at the .10 level. Text 2 showed significant differences between the

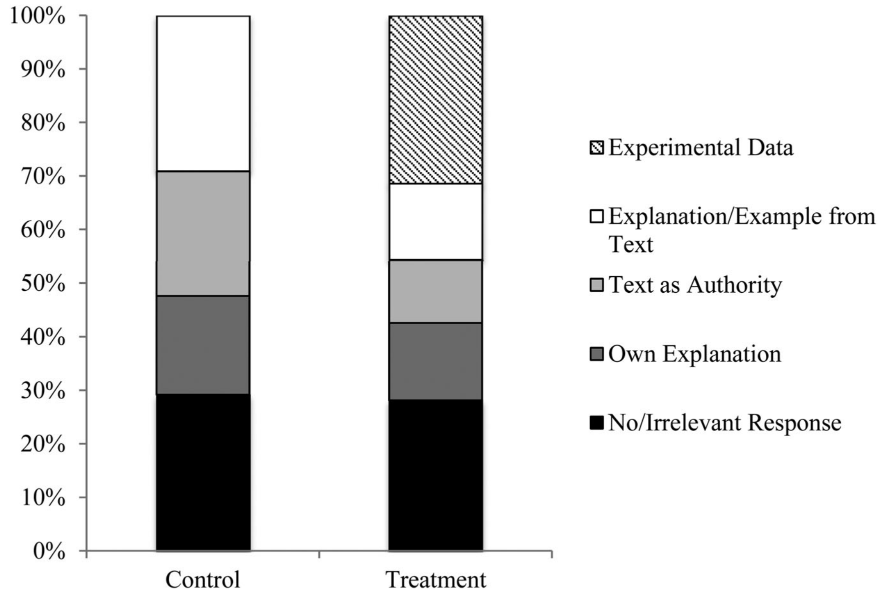

Figure 1. Percentage of students' responses by condition across the three replications indicating their identification of evidence for stated claims in the text.

treatment and control conditions for the conceptual level transfer questions, $F(1) = 14.22$, $p < .001$. Students in the treatment condition scored higher ($M = 3.04$, SE = 0.133) on the conceptual questions than students in the control condition ($M = 2.33$, SE = 0.133). A consistent trend for both the conceptual questions and the design-an-experiment transfer task indicates that students in the top half of the ELA median split outscored students in the lower half of the ELA median split.

Activation of epistemological resources

Evidence for claims

Across the six opportunities in which students were asked to identify evidence for claims in the post-test questions – two for each text replication – approximately 30% of the responses from students in both the control and treatment conditions provided either no response or irrelevant responses to the prompt (Figure 1). *Post-hoc* analyses, using a Bonferonni adjustment for multiple comparisons, showed that students in the treatment condition were more likely to cite experimental data as evidence for the text's claims at a higher rate, approximately 30%, than students in the control condition, across all six instances in which the question was asked. Further *post-hoc* analyses showed that in all six instances in which the question was asked, students in the control condition responded with the 'text as authority' category at a significantly higher percentage than students in the treatment condition. In five of the six instances, students in the control condition provided an 'explanation/example from the text' as evidence for the claim at a significantly higher rate than students in the treatment condition ($p < .05$).

Citing evidence in explanations

Students in the treatment and control conditions differed in their use of evidence when asked for explanations within a transfer task. Importantly, students were asked to explain what would occur in the situation, not justify or provide evidence as part of the explanation. Regardless of condition, students overwhelmingly and expectedly provided responses that were categorised as causal explanations without supporting evidence. Eighty-two per cent of student responses in the control group and 68% of student responses in the treatment group provided a causal explanation that did not include evidence to support the explanation (Figure 2). However, across four of the six questions, *post-hoc* analyses of the split-plot ANOVA indicated a significant interaction between experimental condition and category, each time resulting from a higher rate of students in the treatment group using experimental evidence as part of their explanation ($p < .05$). Despite this statistical difference, less than 10% of students in the treatment condition provided an explanation with reference to experimental data.

Discussion

Results from this study are discussed in the order of the stated research questions and in light of frameworks for reading comprehension, learning from texts, and the resource-allocation model of personal epistemologies. I conclude with the limitations of this study as well as areas for future research in alternative science texts.

Interest

Previous research has shown that interest is an important component for learning from texts as it can mediate attention allocation and help create links to prior knowledge (Schiefele, 1992). No main effect for interest by condition was detected – mean scores on the interest survey and the numbers of questions asked were statistically similar for both

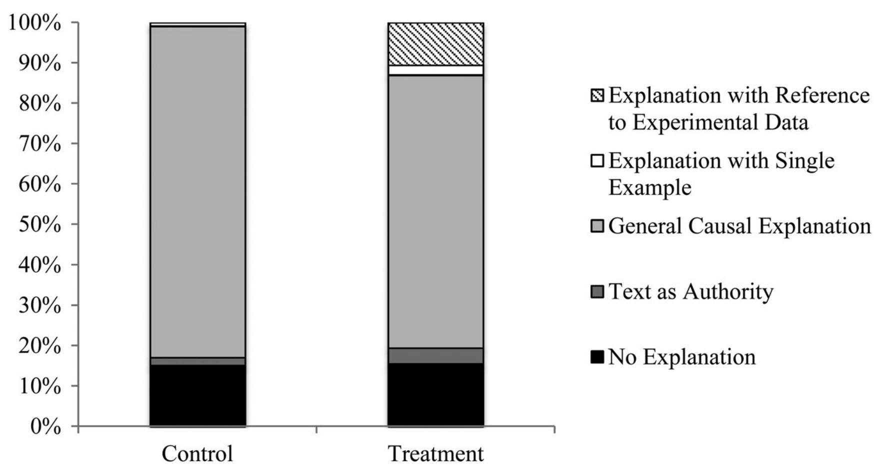

Figure 2. Percentage of students' explanations that used different forms of justification by condition and replication.

EC and TT conditions across all three texts. This suggests that although the EC texts were constructed in a more narrative style with descriptions of the scientists, their questions, and their experiments, this structural change had no differential effect on student interest when compared to TT accounts. In light of our theoretical framework for reading in which interest mediates attention allocation, these results suggest that student attention allocation likely was not impacted by the more epistemically considerate structure of the EC texts.

It should be noted that while the main effect for interest was not significant, results from Text 1 showed a significant interaction between condition and ELA prior achievement. Students with higher prior ELA achievement, that is, stronger readers, who read the EC account, reported a higher average on the interest survey. Text 1, which focused on cell signalling, was also the most difficult of the three text pairs with a reading level of just over tenth grade. The experiment described in the EC account for Text 1 focused on the transplantation of cells from a frog embryo to another frog embryo that resulted in a two-headed tadpole. Previous literature has shown that visceral topics can generate more interest among readers (Schank, 1979); it is possible that the presence of the startling two-headed tadpole caught the attention of readers, but the semantic complexity of the tenth grade reading level along with the complexity of the transplantation experiment made this phenomenon more accessible to the stronger readers.

While the experiment replications suggest no main effect of the text type on interest and only a single interaction by condition and reading level, it is important to think about these results in light of similar students who provided cognitive interviews about their interest levels while reading both text versions (Kloser, 2013). Students who performed think-alouds using these texts were from the same student population, but not involved in the experiment. In the think-alouds study, students were asked to read the control and treatment texts consecutively; the text order was counterbalanced across the sample of 24 students. After reading both texts, students were asked to make a direct comparison of which text they found more interesting. Eighteen of the 24 students chose the EC text and 6 students chose the TT account. These findings suggest that on a global scale, students may not have found the EC texts in this randomised experiment interesting compared to all other interactions in their lives. But, if asked to compare the EC and TT texts to texts that they normally read in science class, perceptible differences might exist. Future experiments with large samples of students could investigate both quantitatively and qualitatively how the interest and attention allocation are related to the different text types.

Comprehension and learning

Student results on the cognitive measures – comprehension and learning – raise questions about the ways in which students semantically analyse and generate mental representations about the texts' content. The lack of any significant main effects of text condition on comprehension may suggest that the elements of the EC texts did not help students decode the semantic structure or create valid mental representations of the texts' ideas any better than the TT account. On one hand, the lack of an impact on comprehension by condition is surprising as the narrative style of the EC accounts and the inclusion of evidence to support the texts' claims theoretically would provide readers more supports for understanding the text as well as creating a mental representation of the text's content. It might even be predicted that these elements might benefit struggling readers even more, yet no such interaction was found in the data.

Measures of comprehension in this study focused on students' ability to reproduce the main idea about the texts. These main ideas focused on fairly simple biology concepts. Given that traditional texts generally use a direct lexicon that states explicitly 'what is', it is possible that any benefits gained by the narrative structure or the explanatory data of the EC texts were countered by this style's less direct approach. Furthermore, by high school, students have engaged in eight to ten years of TT reading and have likely developed a schema for this type of prose. Students generally know where to find the main idea; they are familiar with sentence structure and recognise the expository format. Existing schema for TTs likely aids most students in the semantic analysis of the text. Evidence from the think-alouds study using these same texts supports the possibility of a schema bias (Kloser, 2013). Many students in the think-alouds struggled, at times, in navigating the EC account because they lacked familiarity with how the ideas were presented. Thus, the negligible differences in comprehension between students who read the EC and TT versions should not necessarily be interpreted as there being no effect of text type, but that further studies need to be conducted with students after they have developed a schema for reading EC type accounts.

Interestingly, although comprehension is a main mediating factor of learning from texts and no difference was found between the two conditions, significant differences were found for some of the transfer tasks. Text 2 showed statistically higher outcomes on the conceptual transfer questions for students reading the EC accounts. And for the experimental design transfer task, students reading the EC accounts for Text 1 scored higher than those reading the TT account while Texts 2 and 3 were approaching significance ($p < .10$). These findings suggest that elements of the EC texts, especially the description of scientific practices used to collect data for the EC texts' claims, may be useful in helping students go beyond rote memorisation. The EC texts provided models of sophisticated scientific practice that may have influenced how students viewed the controlling of variables, the collection of data, and what results would support or negate a particular claim. It is possible that the visual representations of the supporting data helped improve coherence among the ideas in the prose (Levin & Mayer, 1993). In short, the students' mental representations of the concepts were richer and more useful when asked near-transfer tasks.

Although students reading EC texts had equivalent comprehension scores as students reading TT accounts, results suggest that what ideas students reading EC texts do comprehend, they are able to better apply these ideas to new contexts and to conceptual situations. This finding has implications for how teachers address new learning standards. New standards in both science and ELA in the United States emphasise that texts serve a purpose beyond comprehension and rote memorisation. Whether in science labs or in science classrooms, obtaining information from texts has the further purpose of using the information and evidence to develop conceptual understanding, construct arguments, and to form causal explanations. Comprehension is a necessary, but not sufficient element of reading to learn in science. While not all values were statistically significant, trends across the multiple replications suggest that elements of the EC accounts help students use the information in a more cognitively complex way. More research and further replications need to occur to see whether the variation in results across text type was due to actual differences within and between the EC and TT versions or whether other factors were responsible.

The fact that significant results varied across text topic may underscore that topic and content play an important role in how students engage texts. Particular topics may be more relatable to students' interest and their prior biological knowledge. Although randomisation assumingly balances unobserved characteristics, like prior knowledge, across conditions, how students' prior knowledge interacts with their reading experience and the generation of a mental representation of the texts' messages is unknown from this data. The qualitative think-alouds run in tandem with this experiment, and provide some supporting evidence that elements unique to EC texts may better facilitate conceptual understanding and transfer. As reported in Kloser (2013), students were more likely to interrupt their reading and think aloud to construct meaning for visual representations in the EC version than for the TT version. For example, 46% of the interruptions with the EC texts were used to construct meaning while only 28% of the interruptions with the TT version were used to construct meaning. For the latter text, the majority of interruptions (44%) involved the description visual representations, a form of comprehension, rather than the generation of inferences or the application of new ideas.

Epistemic sophistication

Results about students' epistemic commitments when reading the EC or TT accounts have both positive and negative implications for students' development of more sophisticated personal epistemologies. On the positive side, Figure 1 shows that when students are given opportunities to read texts that have evidence to support claims, many students cite this empirical evidence as justification for claims in the text. Approximately 30% of students cited empirical evidence as a justification for claims in the EC version. On the negative side, while empirical evidence was the most common response among students reading EC accounts, a majority of students did not cite empirical evidence as a source for justification and a direct comparison with students in the TT version is not possible since by definition the TT versions did not have empirical data that students could cite. What is interesting, though, is that the number of students who read EC accounts who cited the textbook as the authority for the claims was significantly less than students who read the traditional text. Thus, although not all students reading the EC accounts exhibited a more sophisticated epistemic stance, when presented with empirical support far fewer cited that their trust in the text's claims came from the authoritative nature of the textbook. This supports the resource-allocation theory of personal epistemologies. Students have the fine-grained resources to make sophisticated epistemic judgements, but the right stimuli and external resources must be present in order for these commitments to be activated. When reading the TT accounts, these epistemic resources do not exist and in their place is an expository, authoritative account. In the EC condition, students have access to empirical evidence and many students cited the data as reasons for trusting the text's claims – a more sophisticated epistemological commitment.

As Figure 1 shows, though, the presence of empirical support and epistemic resources is not sufficient for many students. Over two-thirds of students in the EC condition did not recognise this resource. Furthermore, in their own explanations (Figure 2), students who read the EC version actually used little of the evidence embedded within the text to support their explanation. This may not be surprising as students were asked to provide an explanation rather than an argument, but only 10% of the students felt independently compelled to provide some level of evidence for their explanation.

Ultimately, these results suggest that students need to be engaged with texts that provide the epistemic resources for students to exhibit sophisticated views about evidence. Students in the think-aloud study (Kloser, 2013) were statistically more likely to select the EC text as more trustworthy than the TT text, largely because evidence and data were used to support the texts' claims. When given different text types, students are better able to recognise the greater evidentiary support provided in the EC texts and that the expository TT texts lack this type of epistemological backing. Students can show more sophisticated perspectives when given the resources to do so and this study shows that not only hands-on investigations or simulations, but also texts can be a source for students to activate their epistemic resources in ways that are more reflective of the discipline.

Limitations

This study occurred in actual classrooms, but the episodic nature of three replications does not reflect students' longitudinal use of these text types. As texts are used consistently throughout a school year, this study could not speak to how use of the treatment text or time could affect students' outcomes. It is possible that the results would be more pronounced in favour of the treatment texts as students developed a stronger schema for reading them. Or, the results from these three replications may reflect something of a novelty effect that could wear off over time.

Of even more importance, this study chose certain controls that required sacrificing some of the text's authentic classroom context. In order to limit the variation of text use that might occur from individual teachers, the study took place outside of teachers' normal practice. Teachers might use the texts in a variety of ways that have more or less effect on student outcomes. For example, the use of the text with graphic organisers or guiding questions would affect the reading experience. As an individual task, this study design also necessarily ignored the socio-cultural impact of discussing texts and how using either the control or treatment texts in a small or whole group environment could have impacted interest, comprehension, or learning. Future studies should examine how the use of different science text types differs in collaborative settings.

Replications or variations of this study might also use a retrospective interest survey to better capture students' relative levels of interest compared to normal texts that they read in science class. Future surveys might better reflect actual differences in interest between the texts by asking questions like, 'Compared to texts you normally read in science class, how interesting do you find the text you just read?' These types of questions might better replicate the levels of interest found in the think-alouds study when students were asked to directly compare which text they found more interesting.

Conclusions and future research

This study showed promise for the use of EC texts in high school biology classrooms. As an alternative form of texts that better reflect the structure of the discipline, EC texts can provide students engagement with core ideas and the practices of science. Initial results from this randomised study suggest that the EC texts can lead to at least the same levels of comprehension as TT accounts while laying the foundation for greater achievement on higher level conceptual and transfer questions. Furthermore, when students are provided with evidence for claims and epistemic backing, evidence exists that their own epistemologies can better reflect what scientists would consider more sophisticated, albeit still at less than desired rates.

Science education is at an inflection point with the introduction of new national standards in science as well as ELA. Across both frameworks, students are asked to engage rich and complex texts and use evidence to support claims. TTs fail, in most cases, to satisfy the breadth and depth of texts that students are supposed to engage about science. Therefore, alternative texts such as historical narratives, APL, or hybrids like the EC texts described in this study should be studied as they find their way into high school classrooms. One recent study showed the possibility of teachers creating forms of APL that were aligned with the original texts and were used to promote discourse of science and disciplinary literacy (Hollingsworth Koomen, Weaver, Blair, & Oberhauser, 2016). Thus, with professional development and training, the availability of these resources for use within the classroom may increase.

Although this study took place in actual classrooms, future studies should investigate even more authentic use across multiple units and how these texts are embedded within the teachers' instruction. Questions remain about how teachers can best facilitate students' comprehension and application of ideas of these texts. Current technologies, such as eye tracking devices in a lab setting, can provide more information about what textual elements students engage most when reading different scientific texts. Given the investment in and importance of texts and inscriptions in science as a discipline, future research should invest time and energy investigating the importance and impacts of differential texts and inscriptions in science as a classroom subject.

Note

1. Pseudonyms are used for all schools, teachers, and students.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This study was supported by the Stanford Graduate School of Education Dissertation Grant.

References

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, *94*(3), 545–561.

Alexander, P. A., & Jetton, T. L. (1996). The role of importance and interest in the processing of text. *Educational Psychology Review*, *8*(1), 89–121.

American Association for the Advancement of Science. (2005). *High school biology: A benchmarks-based evaluation*. Retrieved April 22, 2010, from http://www.project2061.org/publications/textbook/hsbio/report/about.htm

Anderson, R. C. (1984). Role of the reader's schema in comprehension, learning, and memory. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to read in American schools: Basal readers and content texts* (pp. 243–258). Hillsdale, NJ: Lawrence Erlbaum.

Bergin, D. A. (1999). Influences on classroom interest. *Educational Psychologist*, *34*(2), 87–98.

Brill, G., & Yarden, A. (2003). Learning biology through research papers: A stimulus for question-asking by high-school students. *Cell Biology Education*, *2*(4), 266–274.

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*(5977), 453–456.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The Landscape Model of reading: Inferences and the online construction of memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (1st ed., pp. 71–98). Mahwah, NJ: Lawrence Erlbaum.

Brush, S. G. (1974). Should the history of science be rated X? The way scientists behave (according to historians) might not be a good model for students. *Science*, *183*(4130), 1164–1172.

Calfee, R. C., & Curley, R. G. (1984). Structures of prose in the content areas. In J. Flood (Eds.), *Understanding reading comprehension* (pp. 161–181). Newark, DE: International Reading Association.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi – Experimental designs*. Boston, MA: Houghton Mifflin.

Chambliss, M. J. (2002). The characteristics of well-designed science textbooks. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science textbooks* (pp. 51–72). Mahwah, NJ: Lawrence Erlbaum.

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

Chiappetta, E. L., & Fillman, D. A. (2007). Analysis of five high school biology textbooks used in the united states for inclusion of the nature of science. *International Journal of Science Education*, *29* (15), 1847–1868.

Duschl, R. A. (1990). *Restructuring science education: The importance of theories and their development*. New York: Teachers College Press.

Elshout-Mohr, M., & van Daalen-Kapteijns, M. (2002). Situated regulation of scientific text processing. In J. Otero, A. C. Graesser, & J. A. León (Eds.), *The psychology of science textbooks* (pp. 223–252). Mahwah, NJ: Lawrence Erlbaums.

Finson, K. D. (2002). Drawing a scientist: What we do and do not know after fifty years of drawings. *School Science and Mathematics*, *102*(7), 335–345.

Garner, R. (1987). Strategies for reading and studying expository text. *Educational Psychologist*, *22* (3), 299–312.

Goldman, S. R., & Bisanz, G. L. (2002). Toward a functional analysis of scientific genres: Implications for understand and learning processes. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science textbooks* (pp. 19–50). Mahwah, NJ: Lawrence Erlbaum.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395.

Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics. *Journal of the Learning Sciences*, *12*(1), 53–90.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127.

Hollingsworth Koomen, M., Weaver, S., Blair, R. B., & Oberhauser, K. A. (2016). Disciplinary literacy in the science classroom: Using adaptive primary literature. *Journal of Research in Science Teaching*, *53*(6), 847–894.

Hurd, P. D. H., Bybee, R. W., Kahle, J. B., & Yager, R. E. (1980). Biology education in secondary schools of the United States. *The American Biology Teacher*, *42*(7), 388–410.

Kloser, M. (2013). Exploring high school biology students; engagement with more and less epistemologically considerate texts. *Journal of Research in Science Teaching*, *50*(10), 1232–1257.
Latour, B. (1986). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.
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.
León, J. A., & Penalba, G. E. (2002). Understanding causality and temporal sequence in scientific discourse. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science textbooks* (pp. 155–178). Mahwah, NJ: Lawrence Erlbaum.
Levin, J. R., & Mayer, R. E. (1993). Understanding illustrations in text. In B. K. Britton, A. Woodward, & M. R. Binkley (Eds.), *Learning from textbooks* (pp. 95–114). Hillsdale, NJ: Lawrence Erlbaum.
Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, *39*(1), 57–68.
Matthews, M. R. (1994). *Science teaching*. New York, NY: Routledge.
Maury, P., Perez, O., & León, J. A. (2002). Predictive inferences in scientific and technological contexts. In J. Ortero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science textbooks* (pp. 199–222). Mahwah, NJ: Lawrence Erlbaum.
McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1–43.
Merzyn, G. (1987). The language of school science. *International Journal of Science Education*, *9*(4), 483–489.
Monk, M., & Osborne, J. (1997). Placing the history and philosophy of science on the curriculum: A model for the development of pedagogy. *Science Education*, *81*(4), 405–424.
Myers, G. (1992). Textbooks and the sociology of scientific knowledge. *English for Specific Purposes*, *11*(1), 3–17.
National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards*. Washington, DC: Author.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
Norris, S. P., Falk, H., Federico-Agraso, M., Jiménez-Aleixandre, M. P., Phillips, L. M., & Yarden, A. (2009). Reading science texts—epistemology, inquiry, authenticity—a rejoinder to Jonathan Osborne. *Research in Science Education*, *39*(3), 405–410.
Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, *87*, 224–240.
Olson, S. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academies Press.
Osborne, J. (2009). The potential of adapted primary literature (APL) for learning: A response. *Research in Science Education*, *39*(3), 397–403.
Otero, J., León, J. A., & Graesser, A. C. (Eds.). (2002). *The psychology of science text comprehension*. Mahwah, NJ: Lawrence Erlbaum.
Phillips, L. M., & Norris, S. P. (2009). Bridging the gap between the language of science and the language of school science through the use of adapted primary literature. *Research in Science Education*, *39*(3), 313–319.
Sandoval, W. A., & Çam, A. (2011). Elementary children's judgments of the epistemic status of sources of justification. *Science Education*, *95*(3), 383–408.
Schank, R. C. (1979). Interestingness: Controlling inferences. *Artificial Intelligence*, *12*, 273–297.
Schiefele, U. (1992). Topic interest and levels of text comprehension. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 151–182). Hillsdale, NJ: Lawrence Erlbaum.
Schwab, J. J. (1962). The teaching of science as inquiry. In J. Schwab & P. Brandwein (Eds.), *The teaching of science* (pp. 1–104). Cambridge, MA: Harvard University Press.

Schwab, J. J. (1978). Education and the structure of the disciplines. In I. Westbury & N. J. Wilkof (Eds.), *Science, curriculum, and the liberal education: Selected essays* (pp. 229–272). Chicago, IL: University of Chicago Press.

Tobin, K., & Gallagher, J. J. (1987). What happens in high school science classrooms? *Journal of Curriculum Studies*, *19*(6), 549–560.

Tulip, D., & Cook, A. (1993). Teacher and student usage of science textbooks. *Research in Science Education*, *23*(1), 302–307.

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom: A study of K-12 mathematics and science education in the United States*. Chapel Hill, NC: Horizon Research. Retrieved from http://www.horizon-research.com/

Wellington, J. J., & Osborne, J. (2001). *Language and literacy in science education*. Buckingham: Open University Press.

Wood, K. (2002). *2000 National survey of science and mathematics education*. Retrieved April 22, 2010, from http://2000survey.horizon-research.com/

Yarden, A., Brill, G., & Falk, H. (2001). Primary literature as a basis for a high-school biology curriculum. *Journal of Biological Education*, *35*(4), 190–195.