International Journal of Science Education

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

An Analysis of Data Activities and Instructional Supports in Middle School Science Textbooks

Bradley J. Morris, Amy M. Masnick, Katie Baker & Angela Junglen

To cite this article: Bradley J. Morris, Amy M. Masnick, Katie Baker & Angela Junglen (2015): An Analysis of Data Activities and Instructional Supports in Middle School Science Textbooks, International Journal of Science Education, DOI: 10.1080/09500693.2015.1101655

To link to this article: http://dx.doi.org/10.1080/09500693.2015.1101655

International Journal of Science Education, 2015
http://dx.doi.org/10.1080/09500693.2015.1101655

Routledge
Taylor & Francis Group

An Analysis of Data Activities and Instructional Supports in Middle School Science Textbooks

Bradley J. Morris[a*], Amy M. Masnick[b], Katie Baker[c] and Angela Junglen[a]
[a]*Department of Educational Psychology, Kent State University, OH, USA;* [b]*Psychology Department, Hofstra University, NY, USA;* [c]*Baker Educational Consulting, East Grand Rapids, MI, USA*

A critical component of science and math education is reasoning with data. Science textbooks are instructional tools that provide opportunities for learning science content (e.g. facts about force and motion) and process skills (e.g. data recording) that support and augment reasoning with data. In addition, the construction and design of textbooks influence the instructional strategies used in the classroom to teach science. An analysis of science textbooks provides a window to examine what students are being taught about data and how they are being taught. We had two objectives for the present study: (1) to examine opportunities for reasoning with data and (2) to examine to what extent these activities are aligned with instructional supports derived from evidence-based learning strategies. We conducted a descriptive study in which we examined how 20 Middle School science textbooks, across 731 activities, presented opportunities for reasoning with data. Our results demonstrate that although half of activities in textbooks included data, very few of these activities provide opportunities to learn how to record, analyze, and interpret data and the activities rarely provided instructional supports based on evidence-based learning strategies. Our analysis suggests that science textbooks provide limited support for reasoning with data.

Keywords: *Data analysis; Science textbooks; Textbook analysis; Instructional supports*

A critical component of science and math education is reasoning with data. Although this emphasis is strong in recent national standards in the United States (Common Core Standards, 2011, measurement & data; Next Generation Science Standards (NGSS), National Research Council, 2012, data and interpretation), it has also

*Corresponding author. Department of Educational Psychology, Kent State University, 405 White Hall, Kent, OH 44242, USA. Email: bmorri20@kent.edu

been a part of recommended science and math teaching long before the current standards were proposed (National Council of Teachers of Mathematics, 2000; National Science Education Standards, National Research Council, 1996). Reasoning with data involves learning mathematical concepts and procedures, acquiring the cultural tools for recording, interpreting, and representing data, and is important for conducting authentic inquiry in science (e.g. hypothesis testing; Koslowski, 1996; Morris, Croker, Masnick, & Zimmerman, 2012; Zimmerman, 2007). Because textbooks are widely used in science instruction, it is important to understand the kinds of data-related content presented in textbooks as well as how this content is presented.

We conducted a descriptive study of data instruction in science textbooks with two objectives. One objective is to analyze how American middle school science textbooks (for students ages 11–14) provide opportunities for reasoning with data. A second objective is to examine to what extent these opportunities for reasoning with data are aligned with empirically demonstrated principles of effective learning (e.g. interleaving, constructing knowledge; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). The rationale for this analysis is that instruction that is aligned with how students learn is more likely to result in the acquisition of durable, flexible knowledge.

Learning Science Content

Teaching science involves much more than simply communicating science facts. It involves teaching students to understand how to use the tools of scientific inquiry, to understand the nature of scientific knowledge and processes (e.g. the role of empirical evidence), and to see themselves as having the potential to contribute to the scientific enterprise (Erduran & Dagher, 2014; Hassard & Dias, 2013; Schweingruber & Fenichel, 2010). Science standards have consistently identified science process skills as critical in achieving the instructional goals outlined above (e.g. NGSS, 2013). These process skills include, but are not limited to, asking testable questions, designing valid experiments, working with data (e.g. recording, analysis, and interpretation), and generating explanations (NGSS, National Research Council, 2012, data and interpretation; National Science Education Standards, National Research Council, 1996).

The phrase *working with data* describes a group of related scientific inquiry processes (NGSS, 2013). We will discuss four of these processes. *Recording data* refers to the process of accurately tabulating the results of investigations (Lehrer & Schauble, 1998; Toth, Klahr, & Chen, 2000). Without instructional support, children often err when recording results or fail to provide sufficient organization to make use of results at a later time (Lehrer & Schauble, 1998). *Analyzing data* is not the process of making sense of data but is instead the process of summarizing and categorizing data (Kerlinger, 1986). For example, inferential statistical analyses allow a comparison between the patterns in the data and the probability of the patterns occurring by chance. Although children rarely conduct such analyses spontaneously, children as young as 8 use intuitive strategies to determine differences between sets of data (Masnick & Morris, 2008). *Interpreting data* is drawing inferences about the phenomenon from data (Kerlinger, 1986). As with analyses, children have intuitions about how to interpret patterns of evidence from data, often using naïve theories to make sense of data. Naïve theories can overpower data interpretation, in that even data that has been recorded and analyzed correctly is sometimes interpreted incorrectly when it is in conflict with prior beliefs (Chinn & Brewer, 1993).

Generating *predictions and explanations* refer to forecasting future occurrences and to understanding why phenomena occur. Accurate predictions and explanations are derived from understanding scientific concepts and their relations, particularly causal relations (Koslowski, 1996; Morris et al., 2012; Zimmerman, 2007). Predictions/explanations also differ in their scope. Specifically, local explanations target specific phenomenon (e.g. why did a ball fall after leaving someone's hand?) while global explanations target classes of related phenomena (e.g. why do unsupported objects fall?).

Generating explanations and predictions supports the construction of knowledge by linking new knowledge to current knowledge, rather than simply encoding information as presented (Chi, 2009; Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Prompts for predictions and self-explanations have been effective in improving learning in physics (Nokes-Malach, VanLehn, Belenky, Lichtenstein, & Cox, 2013), geometry (Aleven & Koedinger, 2002), and logical proofs (Hodds, Alcock, & Inglis, 2014).

Textbooks as Instructional Tools

Textbooks are commonly used as either the primary (46%) or secondary instructional resource (40%) in science classrooms (Kesidou & Roseman, 2002; Mullis, Martin, Foy, & Arora, 2012; Mullis, Martin, Minnich, et al., 2012). As a result, textbooks heavily influence the structure and direction of science lessons. For example 63% of science teachers follow the 'conceptual framework of the [textbook's] content' (Sánchez & Valcarcel, 1999, p. 500). Although following the lead of the text, science teachers report being unsatisfied with the nature of this instruction (Bryce, 2011; Stansfield, 2006).

Students often hold misconceptions about content areas in math and science (Morris et al., 2012). When textbooks inadequately present science concepts, they fail to address many naïve beliefs and misconceptions about science that students hold (Abd-El-Khalick, 2002). Thus, students who arrive in the classroom with erroneous ideas about cause and effect relations in science (e.g. mistakenly believing that heavy objects fall faster than light objects) can have those incorrect beliefs entrenched with ineffective instruction (Kirschner, Sweller, & Clark, 2006).

One reason for the persistence of such misconceptions is that students are often presented with science as a collection of concepts to be memorized rather than as a process for information-seeking and evaluation (Duit & Treagust, 2012; Kuhn, 2010; Minner, Levy, & Century, 2010; Zimmerman, 2007). Science instruction that includes reasoning with data provides opportunities to learn the value of evidence and how to interpret theory in light of evidence (Masnick & Morris, 2008). Reasoning with data is critical for practicing scientists and is foundational for science education

(Koslowski, 1996). Reasoning with data includes integration with science concepts, effectively recording data, using data analysis strategies, data interpretation, and prediction (Kuhn, 2010). When students are asked to work with data in the classroom, they tend to be given highly constrained data and focus on whether or not their conclusions are correct (Toth et al., 2000). Working with data can provide another path for generating new information and can be a way to learn about a causal mechanism for a phenomenon, both of which may help children construct accurate knowledge and reduce misconceptions (Koslowski, 1996).

There is some work examining the links between pedagogical knowledge and activities in science texts. One recent case study of two high school physics textbooks, one from the US and one from Finland, assessed inquiry-learning opportunities suggested in the texts (Park & Lavonen, 2013). They found that in both texts, fewer than half the questions in the text sampled asked students to reflect on laboratory experiences. Further, the majority of questions that weren't about experiments were simple direct information questions. Data from Project 2061, examining middle school science texts, found that the alignment between national standards and text content was very limited (Stern & Roseman, 2004), and that known approaches to effective pedagogy were rarely used (Budiansky, 2001).

Previous textbook analyses have focused mostly on the quality of the text content. However, as noted above, US science education standards (National Research Council, 1996) and math education standards (National Council of Teachers of Mathematics, 2000) have emphasized the value of reasoning with data. To date, there has been no systematic investigation focused on how textbooks developed to teach to these standards present opportunities for reasoning with data.

Evidence-based Learning Strategies

As noted above, textbooks are highly influential in how teachers structure science education. Because of this role, it is important to assess whether textbooks provide instructional supports for reasoning with data. It is important to investigate whether textbooks structure opportunities for learning that are aligned with learning strategies that have empirical evidence to demonstrate their effectiveness. An example of support for learning how to reason with data is providing detailed instructions for the use of procedures (i.e. worked examples) that lead to accurate understanding of scientific phenomena.

Research from educational and developmental psychology has greatly improved our understanding of effective teaching and learning, by documenting the empirical validity of several different approaches. Although research is currently being conducted on these approaches, most have been identified and investigated for decades (e.g. flexible knowledge; Brown, 1990). We focused on four evidence-based instructional strategies.

1. Integration between activity and text (hereafter Integration). Linking new activities to course content is important for learning because content activates prior knowledge (Hewson & Hewson, 1983; Tricot & Sweller, 2014). It is critical to activate task-relevant knowledge; otherwise students may learn incorrect information or fail to learn because of excessive cognitive load imposed in the learning context (Kirschner et al., 2006). There is limited work exploring how well-integrated text and activities are, though in middle school texts the text is often not linked with graphical images in the text (Slough, McTigue, Kim, & Jennings, 2010). Slough et al. found that although about 60% of graphics were right next to or on the page facing the related text, approximately one-third of the time, figures were unconnected with text. Figure captions largely described or identified the figure, but fewer than 10% of the captions suggested ways to engage students with the content of the figure.

2. Practice. Distributed practice (also known as distributed learning), involves spacing practice across time, and is contrasted with massed practice, or relegating practice to a single session (Bahrick, 1979). Distributed practice is associated with better learning, better retention, better transfer, and faster re-learning compared to massed practice, such as cramming for an exam (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Dunlosky et al., 2013). For example, 5–7-year-olds who were given a science lesson that was repeated across 4 days, demonstrated better learning and transfer than children given four repetitions of the lesson in a single day or two repetitions over 2 days (Vlach & Sandhofer, 2012). These findings suggest that texts ideally should consider repeating lessons across chapters or segments, so that students have repeated exposure to the material. Although the recent research shows this effect in elementary science materials, other evidence of the value of spacing in learning in other contexts has been demonstrated since the 1970s including music (Simmons, 2011) and learning history (Carpenter, Pashler, & Cepeda, 2009). Although practice itself is important, the materials on which practice occurs can increase its effectiveness. One such strategy is interleaving or using more than one kind of problem example in a practice set (Rohrer, 2009; Taylor & Rohrer, 2010). For example, instead of having students solve only addition problems, they might solve a mixture of addition, subtraction, multiplication, and division problems. Compared to learning homogenous materials, learning with multiple types of materials takes more time and effort (Shea & Morgan, 1979). However, long-term retention and transfer are superior when students are presented with interleaved materials (Shea & Morgan, 1979; Son & Simon, 2012). For example, 5th and 6th graders showed the strongest gains in learning about fractions when given interleaved practice with different problem types (Rau, Aleven, & Rummel, 2013). Thus, an ideal text would encourage activities that draw on not only the current chapter's materials, but also relevant concepts from earlier parts of the book, providing an interleaved structure for continued reinforcement of concepts throughout the text.

3. Support for solutions. Effective instruction provides students with support for solutions, particularly during initial exposures to problems. Generating a novel solution requires significant working memory capacity, thus students benefit from high levels of instruction support for initial problems to offset this cognitive load (Kirschner et al., 2006). One evidence-based practice that provides such support is the presence of worked examples. Worked examples provide clear and detailed examples so that children have models for the step-by-step instructions necessary to follow new procedures

(Sweller & Cooper, 1985). Because working with data often requires implementing a set of procedures in a prescribed order, worked examples are beneficial for instruction in this domain. Worked examples are effective in promoting learning multi-step processes such as those encountered in algebra problems (Rittle-Johnson, 2006), calculating compound interest (Renkl, Stark, Gruber, & Mandl, 1998), and calculating the size of an image in optics problems (Ward & Sweller, 1990).

4. Supporting Knowledge Construction. Correcting misconceptions. As discussed above, children often come to science classes with misconceptions about phenomena (Morris et al., 2012) or acquire misconceptions during instruction (via inadvertently misleading instructional materials; McNeil, 2008). Students often look for evidence to confirm their naïve beliefs or existing misconceptions about phenomena (Chinn & Malhotra, 2002) and, thus, misconceptions can become entrenched without careful instructional design (Abd-El-Khalick, 2002). Therefore, an ideal textbook will provide activities that seek to help students revise and correct these misconceptions. Appropriate data interpretation can provide information for challenging and potentially correcting misconceptions. Well-designed instructional materials that help students interpret data provide another source of support for constructing accurate scientific knowledge.

In sum, scientific inquiry involves learning about the individual parts of the scientific process such as generating hypotheses, collecting and recording data, analyzing data, and interpreting data. Instruction that provides support using the evidence-based learning strategies outlined above will be likely to improve students' understanding of each of these stages. To this end, we examined US middle school science textbooks to see how the texts' suggested activities match instruction of science processes using evidence-based learning strategies. We were interested in how well-aligned these texts are with both national curricular goals and with good pedagogical practice.

Method

We coded 20 Middle School Science textbooks published after 2004. Ten of the textbooks were General Science Textbooks from five different publishers, which cover a range of science topics and are typically used for the full year of science instruction. Ten textbooks were specialized textbooks devoted to a specific content area (e.g. force and motion). Books were chosen by Amazon.com sales rank. The full list of books coded is in the Appendix.

The General Science texts included 442 activities and the specialized texts included 271 activities. Three undergraduate students coded all 713 activities. To measure inter-rater reliability, 20% of the activities were double coded. There was 94% agreement on codes and all discrepancies were resolved through discussions with the first author.

Raters coded each activity on 15 categories of good pedagogical practice that are consistent with principles of learning with strong empirical support (see Table 1 for categories and conceptual and operational definitions). In each case, raters gave the activity a '1' if it met the criteria and a '0' if it did not. For example, 'The data

Table 1. Coding categories, definitions, operational definitions, and results by textbook type

Category	Definition	Concept measured	Coding criteria	Mean frequency %	Range %
A. Number of Opportunities	How many times are students asked to perform any type of data operation?		Counted total number of activities in which any type of data were presented	50	19–79
B. Integration	1. Does text refer to activity?	Knowledge activation (McNeil, 2008)	Concepts mentioned within 3 pages of activity	4	0–7
To what extent was the activity integrated with the chapter content?	2. Is the data example/ activity discussed in the chapter summary?	Knowledge activation (McNeil, 2008)	The activity or process was mentioned in summary	1	0–2
	3. Is the same formula/ strategy used in another problem later in the book?	Distributed practice (Vlach, Sandhofer & Kornell, 2008); Interleaving examples (Rohrer, 2009)	Another activity used the same strategy/formula	0	
C. Data Recording	1. Step-by-step instructions to use table or chart?	Worked examples (Rittle-Johnson, 2006)	Example included step-by-step instructions	3	0–7
How were students asked to record/display/organize data?	2. Is there an explanation for why students use table or chart?	Levels of knowledge generation (Chi, 2009)	Explanation for why this procedure was used	10	8–12
	3. Information about how table/chart can be used for different problems?	Distributed practice (Vlach et al., 2008); Interleaving examples (Rohrer, 2009)	Information about using this procedure in a different problem context	0	
D. Data Analysis	1. Asked to analyze data with complex math operation (e.g., formula, series of operations)		Activity required the use of a formula or multi-step mathematical procedure	5.5	0–9
How were students asked to analyze numerical data?	2. Asked to do simple math only (add, subtract, multiply, fractions)	Conceptual vs. procedural knowledge (Rittle-Johnson, 2006)	Activity required a specific simple math procedure (e.g., addition)	25.5	12–39

(*Continued*)

Table 1. Continued

Category	Definition	Concept measured	Coding criteria	Mean frequency %	Range %
	3. Asked to estimate		Activity required the student to estimate possible outcomes	2.5	0–17
	4. Asked to measure (weigh, measure with a ruler, plot)		Activity required measurement of any kind	36.5	23–53
	5. Given step-by-step instructions for performing analysis	Worked examples (Rittle-Johnson, 2006)	Activity included step-by-step instructions	18.5	10–29
	6. Given information on when to use this analysis again	Distributed practice (Vlach et al., 2008); Interleaving examples (Rohrer 2009)	Information provided about using this procedure in a different problem context	0	
E. Data Interpretation	1. Asked to informally compare (look at the differences)	Inducing conceptions and misconceptions from examples (McNeil, 2008)	Students were not asked to perform any type of analysis	64	45–73
How were data used to explain the phenomenon or question in the activity?	2. Asked to formally compare results	Levels of knowledge generation (Chi, 2009)	Students were asked to compare results based on analysis of data (i.e., evidence)	15	0–25
	3. Asked to report largest, fastest, most, etc.		Students were asked to report a single outcome (e.g., fastest time)	17	8–24
	4. Asked to suggest alternative explanations or interpretations of the data		Students were asked to suggest alternative to initial hypothesis or to imagine alternative data/outcomes	13.5	0–18
F. Prediction/ Explanation	1. Local Prediction/ Explanation	Benefits of generating explanations and predictions (Chi, 2009)	Prediction was limited to specific hypothesis or phenomenon (e.g., motion of cars)	44	21–52
Were students asked to make predictions from data?	2. Global Prediction/ Explanation	Constructing knowledge (Chi, 2009)	Prediction was linked to larger class of phenomena (e.g., nature of motion)	8	0–12

example/activity was discussed in the chapter summary', was coded as 1 if the concept in the activity was mentioned in the chapter summary and 0 if it was not.

Results

There were no significant differences based on textbook type, so all data are summarized together. We describe the results of our analyses below and present numerical results in Table 1, which includes the mean frequency of activities from each category and the range of results across the set of 20 texts. Of the 713 total activities, 57% of the activities provided some opportunity to work with data. There were large differences in the number of opportunities for individual texts (range 19%–79%). Recall that our coding focused on how textbooks present instruction on four components of processing data. Within each of these components, we coded whether the activity provided instructional support.

Our results indicated that there were few activities in which students were given the opportunity to record data (2.5% of activities). These data recording activities were associated with little instructional support. Specifically, these activities provided few worked examples (i.e. step-by-step examples) for how to record data and few supports for students to construct explanations for why data were recorded in this way.

Data Analysis. Our results indicated that textbooks provided few opportunities for students to conduct formal analyses of data. Instead, most activities either required no analysis or requested informal analyses (e.g. estimation). The most frequent procedures used were simple math operations (e.g. adding, subtracting) or measuring (e.g. weighing), while few activities asked students to use formulas or estimate. One concern with omitting analysis is that it may generate (or entrench) misconceptions about the role of data. There was little instructional support for data analysis activities. Only 3% of data analysis activities provided step-by-step instructions for data analysis (range 0%–7%) and none of the analytical formulas or techniques were used again, providing no opportunity to practice their use. Finally, no information was provided to support the construction of knowledge about why these procedures would be used and when they might be useful in other content presented in the text.

Data Interpretation. Nearly two-thirds of activities asked for informal interpretations (e.g. what do you think?). Approximately 15% of activities did request interpretations based on data analysis and approximately 14% requested alternative explanations. Alternative explanations are particularly important in constructing robust knowledge of phenomena, for example, speculating on causal mechanism (Koslowski, 1996).

Predictions/Explanations. A related question examined whether the activities prompted local predictions/explanations (i.e. related to the specific activity) or global predictions/explanations (i.e. related to a broad theoretical construct or a general class of phenomena). Nearly half of all activities provided prompts for local predictions, while less than 10% provided prompts for global predictions. Specifically, 44% of the prompts for predictions or explanations were directed to the hypothesis/phenomena in the activity and did not link to other related concepts. Only 8% of prompts for predictions/explanations provided links to higher-order concepts. An example of a prompt that did provide such a link might be an activity on inheritance of traits in plants that includes a prompt to compare or contrast seemingly dissimilar phenomena that share underlying causal mechanisms (e.g. inheritance in plants and animals).

Discussion

This study provides the first systematic examination of how science textbooks provide structure for reasoning with data. The results demonstrate that about half of the activities in science textbooks we examined provided an opportunity for reasoning with data. However, these activities provided few opportunities to learn how to process data (e.g. record, analyze, and interpret data). Further, the activities were rarely presented with instructional support based on well-established learning principles such as interleaved examples or knowledge construction. For example, none of the activities provided the opportunity for students to use a specific data analysis procedure (e.g. comparing means) more than once, which limits a student's opportunity to practice such techniques. Because the activities rarely prompted interpretations based on formal analyses of data, data activities provided few opportunities to address science misconceptions. For example, many activities with data simply ask students to interpret data informally (e.g. 'What do you think happened?').

In conclusion, our analysis of science textbooks suggests that textbooks provide little support for reasoning with data. Our analysis suggests many promising ideas for structuring science classrooms to support reasoning with data. It is important to use data to help students address their naïve beliefs and misconceptions. Specifically, texts and teachers ideally should use data as a tool for supporting inquiry rather than simply as tool for demonstrating science concepts. Given the interest in applying empirically demonstrated learning principles in the classroom, it is also worth considering applying the same principles to the design of textbooks. Finally, using data as evidence to support interpretations of phenomena is a vital goal of science education. These steps in using data in science instruction are more in line with pedagogical evidence than current practices demonstrated in our analyses. If textbooks and in turn teachers, adjusted their focus to align their science teaching in this way, it is likely student understanding of the role of data in science would improve.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abd-El-Khalick, F. (2002). Rutherford's enlarged: A content-embedded activity to teach about nature of science. *Physics Education*, *37*(1), 64–68.

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive Tutor. *Cognitive Science*, *26*(2), 147–179.

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, *108*(3), 296–308.

Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, *14*(1), 107–133.

Bryce, N. (2011). Meeting the reading challenges of science textbooks in the primary grades. *The Reading Teacher*, *64*(7), 474–485.

Budiansky, S. (2001, February). The trouble with textbooks. Prism, 24–27. Retrieved from http://www.project2061.org/publications/articles/articles/asee.htm

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of US history facts. *Applied Cognitive Psychology*, *23*(6), 760–771.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*(1), 1–49.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*(2), 175–218.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58.

Duit, R. H., & Treagust, D. F. (2012). Conceptual change: Still a powerful framework for improving the practice of science instruction. In C. D. T. Kim & M. Kim (Eds.), *Issues and challenges in science education research* (pp. 43–54). Dordrecht, The Netherlands: Springer.

Erduran, S., & Dagher, Z. (2014). *Reconceptualizing the nature of science in science education: Scientific knowledge, practices and other family categories*. Dordrecht: Springer.

Hassard, J., & Dias, M. (2013). *The art of teaching science: Inquiry and innovation in middle school and high school*. New York: Routledge.

Hewson, M. G., & Hewson, P. W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching*, *20*(8), 731–743.

Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, *45*(1), 62–101.

Kerlinger, F. N. (1986). *Foundations of behavioral research*. New York: Holt, Rinehart and Winston, Inc., American Problem Series.

Kesidou, S., & Roseman, J. E. (2002). How well do middle school science programs measure up? Findings from Project 2061's curriculum review. *Journal of Research in Science Teaching*, *39* (6), 522–549.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Boston: The MIT Press.

Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, *94*(5), 810–824.

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

Masnick, A. M., & Morris, B. J. (2008). Investigating the development of data evaluation: The role of data characteristics. *Child development*, *79*(4), 1032–1048.

McNeil, N. M. (2008). Limitations to teaching children 2+2=4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development*, *79*(5), 1524–1537.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496.

Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In H. Kloos, B. J. Morris, & J. L. Amaral (Eds.), *Current topics in children's learning and cognition* (pp. 61–82). Rijeka, Croatia: InTech.

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics* Amsterdam: International Association for the Evaluation of Educational Achievement.

Mullis, I. V., Martin, M. O., Minnich, C. A., Stanco, G. M., Arora, A., Centurino, V. A., & Castle, C. E. (Eds). (2012). *TIMSS 2011 encyclopedia: Education policy and curriculum in mathematics and science* (Vol. 1). Amsterdam: International Association for the Evaluation of Educational Achievement.

NGSS Lead States. 2013. *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics (NCTM).

National Research Council. (1996). *National science education standards. National Committee on Science Education Standards and Assessment*. Washington, DC: The National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education*, *28*(4), 1237–1263.

Park, D., & Lavonen, J. (2013). An analysis of standards-based high school physics textbooks of Finland and the United States. In M.S. Khine (Ed.), *Critical analysis of science textbooks: Evaluating instructional effectiveness* (pp. 219–238). Dordrecht, The Netherlands: Springer.

Rau, M. A., Aleven, V., & Rummel, N. (2013). Interleaved practice in multi-dimensional learning tasks: Which dimension should we interleave? *Learning and Instruction*, *23*, 98–114. doi:10.1016/j.learninstruc.2012.07.003

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology*, *23*(1), 90–108.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1–15.

Rohrer, D. (2009). The effects of spacing and mixing practice problems. *Journal for Research in Mathematics Education*, *40*, 4–17.

Sánchez, G., & Valcarcel, M. V. (1999). Science teachers' views and practices in planning for teaching. *Journal of Research in Science Teaching*, *36*(4), 493–513.

Schweingruber, H. A., & Fenichel, M. (2010). *Surrounded by science: Learning science in informal environments*. Washington, DC: National Academies Press.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(2), 179–187.

Simmons, A. L. (2011). Distributed practice and procedural memory consolidation in musicians' skill learning. *Journal of Research in Music Education*, *59*(4), 357–368.

Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth grade science texts. *Reading Psychology*, *31*(3), 301–325. doi:10.1080/02702710903256502

Son, L. K., & Simon, D. A. (2012). Distributed learning: Data, metacognition, and educational implications. *Educational Psychology Review*, *24*(3), 379–399.

Stansfield, W. D. (2006). Textbooks: Expectations vs. reality: The DNA story. *The American Biology Teacher*, *68*(8), 464–469.

Stern, L., & Roseman, J. (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.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59–89.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, *24*(6), 837–848. doi:10.1002/acp.1598

Toth, E., Klahr, D., & Chen, Z. (2000). Bridging research and practice: A cognitively based classroom intervention for teaching experimentation skills to elementary school children. *Cognition and Instruction*, *18*(4), 423–459.

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, *26*(2), 265–283.

Vlach, H. A., & Sandhofer, C. M. (2012). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development*, *83*(4), 1137–1144. doi:10.1111/j.1467-8624.2012.01781.x

Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in young children's memory and category induction. *Cognition*, *109*, 163–167. doi:10.1016/j.cognition.2008.07.013

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, *7*(1), 1–39.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223.

Appendix: List of Coded Textbooks

Title	Year	Publisher
General		
Glencoe Science Level Red	2008	Glencoe
Glencoe Science Level Green	2008	Glencoe
Glencoe Science Level Blue	2008	Glencoe
McGraw Hill: Science A Closer Look (grade 5)	2007	McGraw Hill
McGraw Hill: Science A Closer Look (grade 6)	2008	McGraw Hill
Prentice Hall Science Explorer: Life Science	2005	Prentice Hall
Prentice Hall Science Explorer: Physical Science	2007	Prentice Hall
Prentice Hall Science Explorer: Earth Science	2004	Prentice Hall
Houghton Mifflin Science, Level 6	2008	Houghton Mifflin
Houghton Mifflin Science, Level 5 B	2008	Houghton Mifflin
Specialized		
Glencoe Science Modules: Electricity and Magnetism	2008	Glencoe
Glencoe Science Modules: Life's Structure and Function	2007	Glencoe
Glencoe Science Modules: Ecology	2008	Glencoe
Holt Science & Technology C: Cells, Heredity, and Classification	2004	Houghton Mifflin
Holt Science & Technology F: Inside the Restless Earth	2007	Houghton Mifflin
Holt Science & Technology I: Weather and Climate	2007	Houghton Mifflin
Holt Science & Technology J: Astronomy	2007	Houghton Mifflin
Interactive Science: Cells and Heredity	2009	Prentice Hall
Interactive Science: Earth's Structure	2009	Prentice Hall
Interactive Science: Ecology and the Environment	2009	Prentice Hall