

How Should Learning Be Structured in Inquiry-based Science Instruction?: Investigating the Interplay of 1st- and 2nd-hand Investigations

S. J. Magnusson, A. S. Palincsar, S. Hapgood & A. Lomangino
4002 SEB, Univ. of Michigan, 610 E. Univ. Ave., Ann Arbor, MI 48109-1259
Tel: 734.616.6855, Fax: 734.936.1606
Email: smag@umich.edu

Abstract: Inquiry-based elementary school science curricula present science learning as predominantly 1st-hand investigations where students directly interact with phenomena. Text-based experiences are limited and not generally integrated with 1st-hand experiences. We conceptualize inquiry instruction as the *interplay* of 1st- and 2nd-hand investigations that are text-based. Our study of 2nd-hand investigation utilized texts that are modeled after key elements of the notebook of a scientist. In an experimental study, we investigated the efficacy of four interplay conditions created by counterbalancing type of investigation (1st- or 2nd-hand) and context of investigation (ramp or table) for learning about the role of mass and force on the motion of objects. Participants were 4th grade students from a low SES working-class community. Results of non-parametric statistical tests of pre-post instruction items on a paper-pencil assessment indicated a preferred sequence and mode of investigation of contexts to support learning when 1st- and 2nd-hand investigations are in interplay. We argue that these results suggest it is possible and important to examine learning from instruction at this level of specificity.

The Problem

The national standard that science teaching will be inquiry-based (NRC, 1996) is defined in the standards document as dealing predominantly with “real phenomena . . . where students are given investigations or guided toward fashioning investigations that are demanding but within their capabilities . . . [but] as more complex topics are addressed [and] students cannot return to basic phenomena for every conceptual understanding . . . teachers can take an inquiry approach as they guide students in acquiring and interpreting information from experts or secondary sources such as text-based or multimedia materials”¹ (p. 31). The primacy placed on direct experience with the physical world –which we might call 1st-hand investigation – is reflected in the kit-based materials that are advocated for use in elementary classrooms. These materials primarily contain physical resources for 1st-hand investigation, and even though additions of some text-based resources have been added in the last couple of years, many teachers see text-based materials as antithetical to inquiry-based science instruction (Shymanski, Yore &, Good, 1991), going so far as to characterize use of text as a step backwards (Alonzo & Jones 2003).

While students’ use of physical materials and involvement in actual investigation are an important departure from textbook-based teaching in which targeted scientific knowledge is “delivered” to the children, the view implied by the national standards – that involvement in investigation alone can foster learning, until concepts become too complex – is an unfortunately simplistic view about learning science. Whereas the standards are partly based on the principle that “school science reflects the intellectual and cultural traditions that characterize the practice of contemporary science” (Ibid., p. 19), the focus of materials at the elementary school level bears only superficial resemblance to the practice of science. First, as Kathy Metz points out, many materials assume that the logic of inquiry is beyond children’s capabilities, engaging children instead in practicing “one or more science process skills” (p. 3, Metz, in press). Second, the texts that are included offer few opportunities to use text in the ways that scientists use text (Crawford, Hurd, & Weller, 1996).

Our conceptualization of inquiry-based science instruction is that children engage in 1st-hand investigation involving inquiry about common phenomena in our world, as well as 2nd-hand investigation in which they learn about and evaluate other’s investigations of the same or similar phenomena, as would a scientist

(Magnusson & Palincsar, 1995). This view has led us to devise text that support such 2nd-hand investigation. These texts are modeled after the notebook of a scientist, and are referred to as notebook texts (Palincsar & Magnusson, 2001). The use of notebook texts is assumed to be in interplay with 1st-hand investigation, as would be the case for scientists. Given that curriculum units are not designed in this fashion, the question arises: Which contexts students should investigate 1st-hand and which are best experienced in a 2nd-hand way, and how should these investigations fit together? While we have identified a set of theoretically-based principles that guide our thinking about the answer to this question (Magnusson & Palincsar, in press), the research described in this paper sought to examine the question empirically.

Theoretical Framework

The writings of biologist-and-curriculum theorist, Joseph Schwab (1962) have influenced our thinking in important ways. Drawing upon his own experiences as a scientist, as well as his knowledge of contemporary work in the history and philosophy of science (e.g., Kuhn, 1970), Schwab argued for 1st-hand investigative contexts that provide opportunities for students to understand the uncertainties and difficulties of knowledge production. He advocated for contexts that feature “phenomena which give rise to problems, the circumstances surrounding the acquisition of data for solving these problems, and the difficulties of working with and among these circumstances,” because such a context no longer tells the student “what to do and what to expect” (Schwab, 1962, p. 54, 55).

With respect to 2nd-hand investigation, Schwab advocated the use of original papers “translated” for specific student use. In his words:

Each individual paper poses the problem of discovering its basic parts (problem, data, interpretation, and so on). Each poses the further problem of discerning the relationship among these parts: why the data sought were the appropriate data for the problem; why the data actually acquired depart from the data sought; what principles justify the interpretation of the data. . . . Are the actual data as appropriate as the reporting scientist considers them to be? What additional assumptions, beyond those noted by the author, are involved in [t]his interpretation? (p. 73-74).

In this spirit, we developed “notebook texts” that model key elements of the notebooks of scientists. In our notebook texts, a fictitious scientist named Lesley Park uses her entries to: (a) identify a problem to investigate, (b) think about how to model the problem for investigation, (c) represent the data collected to support its analysis, (d) describe claims she believes she can make from these data, (e) respond to the critical reactions of her colleagues as they weigh the evidence for her claims, and (f) revise her thinking in light of new data. Notebook texts typically begin with the scientist describing a real-world event she noticed and decided to study, followed by a description of how it was studied, and a series of investigations in which data were collected, analyzed, and claims were made. Figures 1 and 2 show excerpts from one of the notebook texts used in

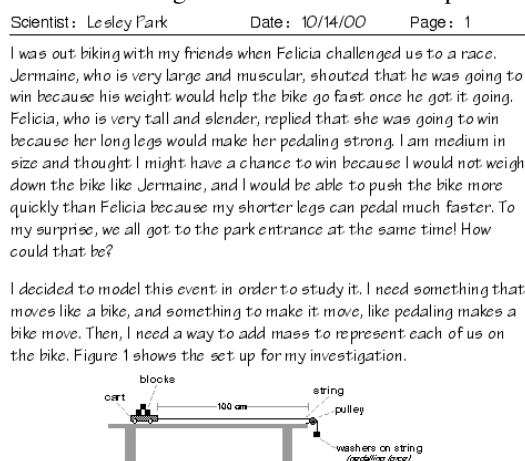


Figure 1 Modeling a person on a bike.

the study reported in this paper.

Figure 1. The first page from Lesley’s notebook regarding the study of horizontal motion.

Table 1 Time of Travel Down a Ramp

Mass on Cart (kg)	Time (seconds)			
	Trial 1	Trial 2	Trial 3	Trial 4
1	2.34	2.38	2.35	2.36
2	2.35	2.38	2.33	2.38
3	2.33	2.37	2.39	2.35

These results do match the outcome of my race with Kendra. So, for some reason, mass does not affect the speed of an object down a ramp. Why would that be? I asked a colleague down the hall, my friend Noah, who is a geologist, to help me think about this question.

I showed Noah my model and my data, and told him my question. He said, "Tell me what you were thinking when you set this up and maybe we can see if you have thought of everything." I said, "I systematically changed the mass. I ran multiple trials. I measured the time. And, I controlled the amount of force." He then asked how I knew that I controlled the amount of force, and I realized that I had not thought about that! I assumed the force was the same because gravity was acting on the cart each time. But is the force of gravity the same if the weight of the object changes?

Figure 2. A page from the Lesley's notebook regarding the study of motion down an inclined plane.

The data collected by Lesley are actual or slightly modified data² that we have collected, and we sometimes feature data from actual scientific investigation of the past (e.g., Newton's study of light through prisms). Notebook texts do not make explicit all of the thinking and decision-making behind what is written, or represented in drawings, partly to resemble an actual notebook where *every* step in the thinking process is not generally spelled out, but also to provide genuine reasons to discuss the information in text, as Schwab encouraged.

Our deliberations about how to design the texts to be used in interplay consider the type of opportunities that they each provide, which we refer to as affordances, as well as the ways in which they might limit learning opportunities, which we refer to as constraints. For example, during second-hand investigations with notebook texts students trace a "single" line of reasoning from question to conclusion, which they are encouraged to critique.³ A single line of reasoning has the advantage of providing students with a simpler conceptual terrain than one featuring multiple lines of reasoning. Moreover, the notebook text presentation places the same information at all of the students' disposal, supporting collaborative reasoning across the classroom community.

In contrast, 1st-hand investigations result in multiple lines of reasoning from question to conclusion across student investigative groups. This can be a more challenging context for students to understand because the conceptual terrain under discussion is less distinct, but it can also motivate students to better concretize and clarify their ideas in order to explore the extent to which their ideas are really the same or different from one another's. Furthermore, 1st-hand investigations provide the opportunity to explore concretely the results of various methods for collecting and analyzing data, which can enrich students' thinking, while at the same time potentially "muddying the waters" for students building new knowledge. Table 1 provides a conceptual comparison of 1st- and 2nd-hand modes of investigation.

The differences in opportunities between these modes of investigation, as well as the trade-offs *within* a mode of investigation, raise the question of how to ascertain when student learning is better supported via work in the more conceptually-varied context of 1st-hand investigation, and when is it better for students to engage in learning via the more streamlined, but "once-removed" context of 2nd-hand investigation. This is the question that motivated the research reported in this paper.

Table 1. Comparison of the attributes of different modes of scientific inquiry.

	Affordances	Constraints
1 st -hand	Direct experiences can be powerful in concretizing scientific relationships describing the physical world.	Variations in the data (due to the complexities of the real world and the many possible sources of error) increase the challenge of seeing patterns in the data. Students may make sense of their experiences in quite different ways from scientists.
	Direct experiences in which one manipulates the physical world, are powerful means for trying out and testing one's thinking.	The social and physical demands of 1 st -hand investigations (e.g., coordinating thinking and activity within a group, coordinating an array of materials) leave little room for students to focus conceptually, requiring additional time for conceptual invention to make meaning of what occurred. Students may lose sight of the targeted question and become more engaged in pursuing their own questions.
1 st -hand (cont.)	Collaborating to produce knowledge claims is an important part of scientific activity.	Children's independent inquiry is not automatically guided by the cultural values, beliefs, norms, and conventions of the scientific community (e.g., need for adequate evidence, role of disconfirming evidence in revising thinking). Thus, students' claims might be quite contrary to the claims developed by scientists, sometimes across years of study.
2 nd -hand	A common set of pertinent information for the doing of science – question, method, data, knowledge claims – is presented to all children.	Interpretation of the information represented in static terms is required, and children's interpretations in the face of static presentation may be erroneous. The static nature of the information in the text may constrain children's abilities to employ the type of reasoning illustrated, when they inquire on their own.
	The processes of thinking that produce scientific knowledge are "laid bare," serving as a model for one's own thinking during scientific investigation.	The process of scientific reasoning is embedded within a context and particular conceptual ideas; thus, is it not transparent, and teacher guidance is required to help students identify and evaluate the scientific reasoning and decision making modeled in the text.

Methodology and Research Design

Our research called for an experimental design with multiple conditions that were counterbalanced for order. We selected to focus on the topic of motion for three reasons: (1) it is commonly studied in the elementary school at the level we were working (in fact, it is now specified by the state of Michigan as a topic for the grade at which we conducted this study), (2) it is an easy context to conduct meaningful investigation of the physical world (i.e., does not require complex procedures or highly specialized or sophisticated tools), and (3) there are relatively easy relationships that can be identified from the study of this topic (e.g., force-motion relationships, mass-motion relationships) as well as more complex relationships (force-mass-motion relationship).

We chose to focus the investigation on the study of motion in two investigative contexts – horizontal motion across a table versus motion down a ramp. Materials for these two contexts were designed such that either context could be investigated in a 1st- or 2nd-hand way. This permitted us to investigate learning across conditions in a 2 x 2 design. Both the topic of investigation and the investigation type were counter-balanced, creating four conditions (Figure 3). Each investigation took five, one-hour sessions, and were taught by the co-PIs of the project (Palincsar and Magnusson). Both contexts for learning about motion were designed to promote the development of student understanding regarding the influence of force and mass on the motion of objects, albeit only in a straight line trajectory.

Participants came from two, intact 4th grade classes from a district with an urban profile, with 51% of the students in the school being African American, and 58.5% receiving free or reduced-cost lunch. Students in each class (n≈24) were matched on a combined score from an assessment of prior knowledge about motion and a standardized reading achievement measure, and were split in half for the study, with students randomly assigned to conditions. Twelve students participated in each condition.

	Investigation Type	Conditions		Investigation Type	Conditions	
		1	2		3	4
WEEK 1	1 st -hand	table	ramp	2 nd -hand	table	ramp
WEEK 2	2 nd -hand	ramp	table	1 st -hand	ramp	table

Figure 3. Experimental research conditions.

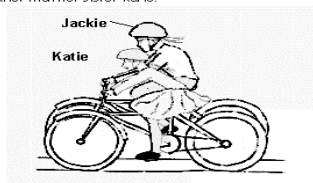
A paper-pencil assessment was administered prior to any instruction, giving us an indication of students' prior knowledge of the topic area, and subsets of that assessment were administered following each investigation – involving only those items pertaining to the context just investigated, giving us an indication of what was learned. The assessment was designed to examine student understanding with respect to scientific content and reasoning regarding the motion of objects, and contained 29 items, approximately half of which concerned motion on a flat surface and half motion down an inclined plane. In terms of scientific *content*, questions addressed the role of force and mass in influencing motion, and the interpretation of time data from changes in those variables presented in tables. Figure 4a shows an example of a science content item. The questions pertaining to scientific *reasoning* concerned the proper use of tools for measuring motion, the use of appropriate procedures for investigating motion, accurate analysis of motion data, and identification of accurate knowledge claims from the results of investigations. Figure 4b illustrates a scientific reasoning item regarding motion.⁴

Findings

Figure 5 shows mean score changes for students by condition, according to type of item – *scientific content* versus *scientific reasoning* – for the two contexts in which students studied motion: across a *table* (horizontal plane) and down *ramp* (inclined plane). Results in term of these subsets of items on the assessment raise the question of whether students' prior knowledge across conditions was equivalent.⁵ Statistical comparisons showed that they were, with the exception of *reasoning* about motion across a *table*: students in Condition 2 had statistically significantly less prior knowledge than students in the other conditions.

One important observation from these results is that there was no one condition in which student performance was most advantaged across both contexts and with respect to the development of scientific content and reasoning. Instead, there was a complex interaction of learning by condition. Furthermore, and unexpected result was that in each condition, there was a decrease in student performance for one of the subtests. We hypothesize that these decreases result from interference effects due to the short time span of the instruction. Specifically, while the variables of mass and force influence motion in both contexts, the influence of mass is not salient in the ramp context; students had to be guided to measure force in relation to mass differences, but indirectly. (stretch of a spring attached to the cart at the top of the ramp). This apparent difference might have interfered with the development of desired knowledge due to insufficient instructional time to fully discuss this issue.

The drawing shows Jackie and her little sister Katie on their bikes. Jackie is much heavier than her sister Katie.



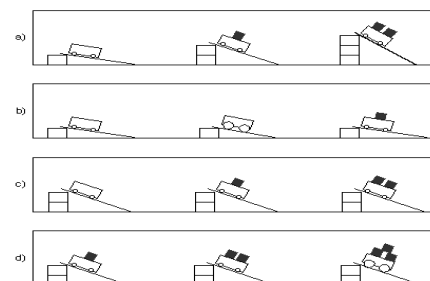
9. If they race their bikes along a flat sidewalk, who will win? Circle the best answer.

- Jackie will win because she can pedal with more force and go faster.
- Katie will win because Jackie's greater mass will slow Jackie down.
- It depends on how much force Jackie can pedal with to make up for her larger mass.
- They will tie because their different masses and different pedaling force will have opposite effects.

(a)

Abdul had some carts and some blocks. The blocks were all the same mass. He wanted to test the idea that *A heavier cart goes down a ramp faster.*

10. Which set-up should he use to test this idea? Circle the best answer.



(b)

Figure 4. Items assessing student science content (a) and reasoning (b) knowledge about motion.

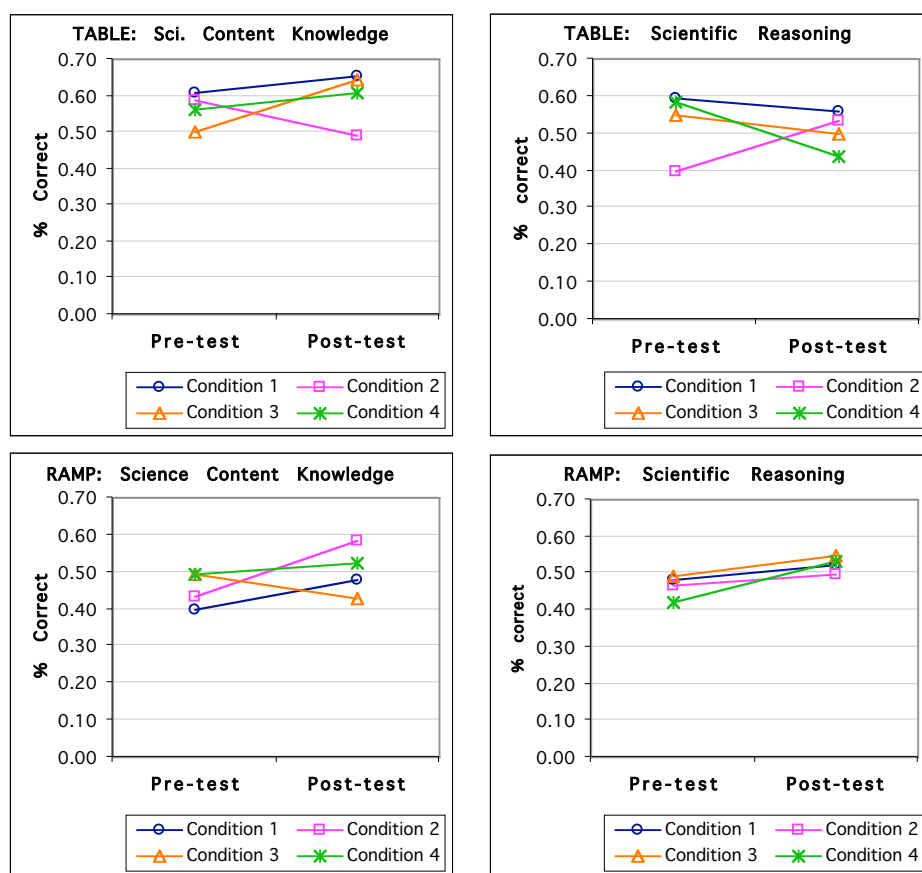


Figure 5. Changes in student understanding about motion across a table (top) and down a ramp (bottom).

Our analysis of these results sought to determine whether we could tease out the effects of the particular elements of the instructional design; that is, the **order** of contexts for learning about motion (table first, ramp second versus ramp first table second), and the **mode** of investigation in which students learned about motion (1st- vs. 2nd-hand with the notebook texts). We wondered whether students' initial investigation of motion, regardless of whether it was about motion across a table or down a ramp, or 1st- or 2nd-hand, advantaged their learning; and we wondered whether students' learning via 1st-hand investigation or 2nd-hand investigation was advantageous, regardless of whether it was their initial or final study of motion. Due to our small sample size, we utilized the non-parametric Kruskal-Wallis test – a one-way analysis of variance for independent samples – to examine these questions. Table 2 shows the results of these comparisons.

Table 2. Significance scores for comparison by condition of results from particular elements of instruction

Element of Instruction	Learning about Sci. Content	Learning about Sci. Reasoning
Initial Investigation	.657	.089
Final Investigation	.089	.004
1 st -hand Investigation	.228	.127
2 nd -hand Investigation	.122	.060

These results indicate that there were no general differences across condition by mode of investigation alone, but students were advantaged in learning about scientific reasoning in their final investigation. Descriptive statistics indicate that increased understanding resulted from Conditions 1 and 2, and these were both cases in which the investigation was 2nd-hand. Conditions 3 and 4 led to decreased understanding. We think this result provides some evidence that 2nd-hand investigations with our notebook

texts advantaged student development of scientific reasoning, but only in contexts where their use followed 1st-hand investigation.

A second set of analyses compared changes in learning by investigation characteristics; that is, whether student learning differed in their initial *versus* their final investigation, and whether it differed when learning via 1st-hand *versus* 2nd-hand investigation. Table 3 shows the Kruskal-Wallis test results of these comparisons.

Table 3. Significance scores for comparison by condition of results from particular characteristics of investigation

Investigation Characteristics	Learning about Sci. Content	Learning about Sci. Reasoning
Final VERSUS Initial Investigation	.163	.002
2 nd -hand VERSUS 1 st -hand Investigation	.036	.011

These results indicate that differences in learning about scientific *reasoning* depended upon the **amount** of experience, as well as the **mode** of investigation experienced. With respect to amount, when 2nd-hand investigations with our notebook text followed 1st-hand investigation, students showed more growth in scientific *reasoning*. With respect to the mode of investigation, the greatest positive change was in Condition 4 and the greatest negative change was in Condition 3. These conditions were opposite in the contexts assigned to 1st- versus 2nd-hand investigation, but both started with 2nd-hand investigations. In that circumstance, it was an advantage to start with a study of the ramp context.

At a more conservative level of significance, differences learning about scientific *content* were statistically significant, but only relative to the type of investigation. Conditions 1 and 3 resulted in increased understanding, and Conditions 2 and 4 resulted in a decrease in understanding. The contrast in results for the extreme cases – Conditions 2 and 3 – are striking because the primary difference was whether the table or the ramp context was studied first. The greatest positive change was for the table context followed by the ramp context (Condition 3). These results indicate that students benefit more if they study motion on a horizontal plane before studying motion down an inclined plane.

Significance

The complexity of the interactions in our results suggest that more research is needed in this vein for us to make broad generalizations. Despite: (1) the relative brevity of instruction (two weeks), (2) the children's inexperience with inquiry-based science instruction, and (3) the relatively small sample sizes, there were statistically significant results that we think make important contributions to our thinking about instruction in science. First, we think the evidence suggesting an advantage to learning from contexts in a particular order is sufficiently strong to suggest that this is an important parameter to examine in research on instruction in science. Second, we think the evidence suggesting some advantages to conducting investigations with notebook texts is sufficiently strong to view it as additional evidence of the benefits of having notebook texts as a part of inquiry-based science instruction. Third, we think these results are strong enough to make an argument for the study of motion involving the use of notebook texts, in the following way: beginning with 1st-hand investigation on a horizontal plane followed by 2nd-hand investigation of motion down a ramp using a notebook text, which we also think is defensible purely on conceptual grounds due to the “hidden” element of force in a ramp context. The fact that our results are consistent with what we would argue for on theoretical grounds suggests a fourth significant point: it is possible to empirically test principles of curriculum and instruction, even with short instructional interventions. We believe this is an important outcome considering the call for teachers to use “best” practices as identified from research. Many of the materials teachers use in science instruction have not been subjected to empirical studies that so closely examine learning with respect to the sequence and/or mode of activity employed. We believe that our findings suggest that such work is needed. Moreover, despite the challenges of conducting rigorous experimental research (Mosteller & Boruch, 2002), our results suggest that meaningful results are possible when conducting this research, even from short instructional interventions. Finally, we also submit that studies of this nature should be replicated, particularly with respect to different topic areas. Our next study will be conducted in a similar vein, and will utilize the findings from this study by examining learning from

the specified interplay condition compared to learning via 1st-hand only and 2nd-hand only investigative experiences.

Endnotes

- (1) The standards say that the processes by which knowledge presented was acquired “should be identified and note sources as “authoritative and accepted within the scientific community” (NRC, 1996, p. 31).
- (2) There are a considerable number of issues that we can discuss regarding our decisions about what data to feature, how much, and to what level of precision, but that is beyond the scope of this paper. Perhaps what is important to know is that our intent to feature data that are as close to what would actually be observed for the specific context. Modifications are sometimes made to accommodate perceived student imitations in knowledge, such as reporting data to the tenths, rather than hundredths place, as occurs in our electricity texts regarding the measurement of current. In the case of the motion text, despite our knowledge that students would be unfamiliar with reading values to the hundredths place, the instruction (which was modified from the *Science and Technology for Children* unit entitled Motion and Design) called for students to observe time with stopwatches that measured time to the hundredths of a second. So, we kept the time of data as we collected it.
- (3) We’ve placed the word single in quotation marks because the fictitious scientist may follow more than one line of reasoning across a notebook text.
- (4) This item is a modified version of a released item that appeared on the TIMSS assessment, which can be accessed via: <http://timss.bc.edu/timss1995i/Items.html>
- (5) Students’ prior knowledge as indicated by total scores for the whole assessment or totals for all the content or all the reasoning items also showed no statistically-significant differences across students in the conditions.

References

- Alonzo, A., Jones, M., Aschbacher, P., & Roth, E. (2003). *Albatross, Add-On, or Tool: How Teachers Conceptualized the Relationship of the Second Hand Investigation Text to a Hands-On Inquiry Science Curriculum*. Paper presented at the American Educational Research Assoc. conference. Chicago, IL, April.
- Crawford, S. Y., Hurd, J. M., & Weller, A. C. (1996). *From print to electronic: The transformation of scientific communication*. Medford, NJ: Information Today, Inc.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Magnusson, S. J. & Palincsar, A. S. (in press). The application of theory to the design of innovative texts supporting science instruction. In M. Conzas & R. Sternberg (Eds.), *Translating Educational Theory and Research into Practice*. Mahwah, NJ: Lawrence Erlbaum and Associates.
- Magnusson, S. J., & Palincsar, A. S. (1995). Learning environments as a site of science education reform. *Theory into Practice*, 34(1), 1-8.
- Metz, K. E. (in press). Children’s understanding of scientific inquiry: Their thinking as designers of research and interpreters of data. *Cognition and Instruction*.
- Mosteller, F., & Boruch, R. (2002). *Evidence matters: Randomized trials in education research*. Washington, DC: Brookings Institution Press.
- National Research Council (NRC) (1996). *National Science Education Standards*. Washington, D.C.: National Academy Press.
- Palincsar, A. S. & Magnusson, S. J. (2001). The interplay of first-hand and text-based investigations to model and support the development of scientific knowledge and reasoning. In S. Carver & D. Klahr (Eds.), *Cognition and Instruction: Twenty-five years of progress* (pp. 151-193). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schwab, J. J. (1962). The teaching of science as enquiry. In J. Schwab & P. Brandwein (Eds.), *The Teaching of Science* (pp. 1-103). Cambridge, MA: Harvard University Press.
- Shymansky, J. A., Yore, L. D., & Good, R. (1991). Elementary school teachers’ beliefs about and perceptions of elementary school science, science reading, science textbook, and supportive instructional factors. *Journal of Research in Science Teaching*, 28 (5), 437-454.

Acknowledgments

This research was supported by a grant from the National Science Foundation under the program for Research on Learning in Education (REC-0207651). All opinions expressed in this work are the authors’ and do not necessarily represent the view of either the funding agency or the University of Michigan.