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

Shirley J. Magnusson, Annemarie S. Palincsar, Adrienne Lomangino, Susanna Hapgood
The University of Michigan

April 2004


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.
Inquiry-based elementary school science curricula present science learning as predominantly 1\textsuperscript{st}-hand investigations where students directly interact with phenomena. Text-based experiences are limited, despite the fact that text use is common in the actual practice of science and national standards in the United States are partly based on the principle that school science reflect the traditions of contemporary science. In contrast to the current state of educational practice at the elementary school level, we conceptualize science instruction as involving the \textit{interplay} of 1\textsuperscript{st}-hand and text-based experiences, which we call 2\textsuperscript{nd}-hand investigations. In an experimental study, we investigated the efficacy of four interplay conditions, created by counterbalancing type of investigation (1\textsuperscript{st}- or 2\textsuperscript{nd}-hand) and context of investigation (motion down a ramp or across a table) for learning about the influence of mass and force on the motion of objects. Participants were 4\textsuperscript{th} grade students from a low SES working class community with \approx 50\% of the population being African American. Second-hand investigations utilized novel texts we have developed that are modeled after key elements of the notebook of a scientist 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. 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” (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 (Shymansky, 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 the study reported in this paper.

**Figure 1.** Part of Page 1 from Lesley’s notebook regarding the study of motion across a horizontal surface.
The data collected by Lesley are actual or slightly modified data 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.

Table 1

<table>
<thead>
<tr>
<th>Mass on Cart (g)</th>
<th>Time 1 (seconds)</th>
<th>Time 2 (seconds)</th>
<th>Time 3 (seconds)</th>
<th>Time 4 (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.34</td>
<td>2.38</td>
<td>2.35</td>
<td>2.36</td>
</tr>
<tr>
<td>2</td>
<td>2.35</td>
<td>2.38</td>
<td>2.33</td>
<td>2.36</td>
</tr>
<tr>
<td>3</td>
<td>2.33</td>
<td>2.37</td>
<td>2.39</td>
<td>2.36</td>
</tr>
</tbody>
</table>

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 goalkicker, 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 masses. 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. Part of Page 2 from Lesley’s notebook regarding the study of motion down an inclined plane.

The data collected by Lesley are actual or slightly modified data 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. 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

<table>
<thead>
<tr>
<th>Mode</th>
<th>Affordances</th>
<th>Constraints</th>
</tr>
</thead>
</table>
| 1st-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 ways of thinking and testing one’s ideas. | • The social and physical demands of 1st-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. |
|       | 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. |
| 2nd-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 subsequent 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. |
|       | The “single” line of reasoning presented makes it easier to follow the | Reasoning to build scientific knowledge rarely follows a single line; hence, such experiences overly simplify the actual production of knowledge under |
production of scientific knowledge from empirical activity. investigative circumstances, and are insufficient to provide the experiences students need to independently build scientific knowledge.

Methodology and Research Design

Instructional Context

Our research involved an experimental design with multiple conditions resulting from the manipulation of the investigative contexts that students encountered in a 1\textsuperscript{st}-hand versus 2\textsuperscript{nd}-hand way. We selected the topic of motion for this study for three reasons: (1) it is commonly studied in the elementary school at the level we were working\textsuperscript{4}, (2) it is an easy context to conduct meaningful investigation of the physical world; that is, it does not require complex procedures or highly specialized or sophisticated tools, and (3) there are unambiguous relationships that can be identified from the study of this topic: the influence of force on the motion of an object, the influence of mass on the motion of an object, and the combined influence of force and mass on the motion of an object.

We focused investigation on the study of motion in two investigative contexts – linear motion across a table and down a ramp. Materials for these two contexts were designed such that either context could be investigated in a 1\textsuperscript{st}- or 2\textsuperscript{nd}-hand way. This permitted us to investigate learning across conditions in a 2 x 2 design, counter-balancing for the topic of investigation and the investigation type (Figure 3). Each investigation took five, one-hour sessions, and were taught by the co-PIs of the project (Palincsar and Magnuson).

![Figure 3. Experimental research conditions.](image)

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. Figure 4a shows the 1\textsuperscript{st}-hand investigation materials used for the study of linear motion across a table, and Figure 4b shows the 1\textsuperscript{st}-hand investigation materials used for the study of linear motion down a ramp.
Participants

Participants were from two, intact 4th grade classes from a district with an urban profile: 51% of the students in the school were African American, and 58.5% received 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 (discussed next) 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.

The prior knowledge assessment was administered prior to any instruction, and subsets of that assessment were administered following each cycle of investigation, but involving only those items pertaining to the context students’ just investigated. Differences in scores on the post-instruction assessment were assumed to be indicators 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 concerned 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 5a shows an example of a science content item. Questions that pertained to scientific reasoning concerned the proper use of tools for measuring motion, the use of appropriate procedures for investigating motion, accurate analysis of data about the motion of objects, and identification of accurate knowledge claims from the results of investigations. Figure 5b shows an item regarding scientific reasoning about motion.5
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:
   a) Jackie will win because she can pedal with more force and go faster.
   b) Katie will win because Jackie’s greater mass will slow Jackie down.
   c) It depends on how much force Jackie can pedal with to make up for her larger mass.
   d) They will tie because their different masses and different pedaling forces will have opposite effects.

(a) (b)

Figure 5. Items assessing student science content (a) and reasoning (b) knowledge about motion.

Findings

Descriptive Statistics

We first present our findings in terms of simple pre-post assessment mean scores by condition, for each context of investigation (table or ramp), and for each type of assessment item (content or reasoning). Figure 6 shows the mean score changes for students by condition, and the results shown raise the question of whether students’ prior knowledge across conditions was equivalent. Statistical comparisons indicated that the only initial difference that was statistically significant, concerned students’ 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 not 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, an 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 via the stretch of a spring attached to the cart when it was 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.
Our analysis of these results sought to determine whether we could tease out the effects of particular elements of the instructional design. The following results utilized statistical analyses to examine student learning relative to various elements of instruction.

The Influence of Order and Type of Investigation

The first elements of instruction that we examine were 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-hand 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 experienced 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
These results indicate that, generally, there were not differences across condition by order or mode of investigation alone, with the exception of the role of order of investigation on students’ learning about scientific reasoning: students seemed more advantaged in developing scientific reasoning in their final investigation. Descriptive statistics for this element of instruction indicate that increased understanding occurred within Conditions 1 and 2. These conditions were both cases in which the investigation was 2nd-hand. Thus, there is some evidence that encountering a 2nd-hand investigation with our notebook texts following a first-hand investigation advantaged student development of scientific reasoning.

Comparison of the Influence of Contrasting Investigation Characteristics

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Final VERSUS Initial Investigation</td>
<td>.163</td>
<td>.002</td>
</tr>
<tr>
<td>2nd- VERSUS 1st-hand Investigation</td>
<td>.036</td>
<td>.011</td>
</tr>
</tbody>
</table>

These results indicate that learning about science content did not appear to be influenced by contrasting investigation characteristics, but differences in learning about scientific reasoning seemed to be influenced by the amount of experience students has with investigation (final vs. initial), as well as the mode of investigation students experienced (2nd- vs. 1st-hand). With respect to amount of experience and student development of scientific reasoning, students in Conditions 1 and 2 showed the most advancement in understanding, and these were the contexts in which students engaged in 2nd-hand investigations with our notebook text following 1st-hand investigation. With respect to the mode of investigation experienced, 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. This is not surprising in that the ramp context is a more complex context to study in a 1st-hand way because the force moving the object is not directly manipulated, it is indirectly measured, and in fact, it varies as a function of mass, though that is not initially evident to students.

If we consider a more conservative level of significance of 0.05, differences learning about scientific content were statistically significant, but only relative to the mode of investigation. Conditions 1 and 3 resulted in increased understanding (learning in Condition 3 was greater),
and Conditions 2 and 4 resulted in a decrease in understanding (learning in Condition 2 showed more of a decrease). 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. 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 relative to these fine points of instruction, for us to make broad generalizations about how inquiry-based science instruction should occur. In addition, similar research needs to be conducted in other topic areas. Nevertheless, despite: (1) the relative brevity of instruction (two weeks), (2) the children's inexperience with inquiry-based science instruction, and (3) our 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 that there is 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 that there are some advantages to conducting investigations with notebook texts is sufficiently strong to suggest that this is also an important parameter to examine in research on instruction in science, and, we view it as additional evidence of the benefits of having notebook texts as a part of inquiry-based science instruction (other evidence is described in Palincsar & Magnusson, 2001).

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 1\textsuperscript{st}-hand investigation on a horizontal plane followed by 2\textsuperscript{nd}-hand investigation of motion down a ramp using a notebook text. What is striking to us is that we think this arrangement of the instruction is defensible purely on conceptual grounds. For example, considering that force cannot be independently manipulated in a ramp context (although force can be indirectly measured), and that force and mass are interdependent (i.e., increasing the mass also increased the force of gravity on an object), motion down a ramp is quite a complex challenging context to investigate in a 1\textsuperscript{st}-hand way. As a result, we would argue that the “single” line of reasoning presented in a notebook text used for 2\textsuperscript{nd}-hand investigation (see Table 1) provides a simpler and streamlined context for students, which we think would enable them to come to understand the desired scientific relationships.

The fact the our results are consistent with what we would argue for on theoretical grounds suggests a fourth significant point:: our results from this study can be considered evidence that it is possible to empirically test principles of curriculum and instruction regarding inquiry-based instruction in science, even with short instructional interventions. We believe this is an important outcome considering the call for teachers to use “best” practices identified from research. Many of the materials teachers use in science instruction have not been subjected to empirical studies that so closely and specifically examine learning science. 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 want to argue for the need to replicate studies of this nature. Replication in educational research is not common, and we think we should be skeptical about results from a single study. In addition, we would argue that such replication would need to be in the same topic area, and that studies in different topic areas should be considered different studies. We make this claim because different topics provide different learning opportunities for students relative to content and reasoning goals, and differences in the types of constructs students are expected to come to understand may change the nature of the relationships between instructional elements. In essence, we call for a response to the point made by Brophy and Alleman (1991) long ago, that we know very little about instruction. We think we can learn much more about instruction relative to the conceptual issues that we have examined in this study. We will continue to take on this kind of research, and hope that this study encourages others to do likewise.
References


Metz, K. E. (in press). Children’s understanding of scientific inquiry: Their thinking as designers of research and interpreters of data. Cognition and Instruction.


Footnotes

1The standards say that the processes by which knowledge presented was acquired” should be identified and to note the sources as “authoritative and accepted within the scientific community” (NRC, 1996, p. 31).

2There 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 it is 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 data as we collected it.

3We’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 in fact, it is now specified by the state of Michigan as a topic of study for the grade level at which we conducted this study – 4th grade.

4This 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

5Students’ 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.