Examining the Development of Students' Hypotheses and Predictions Across Three Inquiry-oriented Curricula

Pier Sun Ho and Nancy Butler Songer
School of Education
The University of Michigan
psh@umich.edu

Paper presented at the Annual Meeting of the National Association for Research in Science Teaching
San Francisco, CA
April 4, 2006

Introduction

One recent trend in science education reform has been the importance of teaching students to think and reason in complex ways. National reform documents emphasize the need for scientific inquiry in science education – engaging students in the activities and thinking processes similar to those of scientists (National Research Council, 2000). Inquiry science provides students with the opportunity to apply scientific skills, knowledge and reasoning to situations that approximate how scientists would do their own work.

Research has shown that the inquiry approach to science manifests numerous cognitive, motivational and epistemological benefits to students. The Cognition and Technology Group at Vanderbilt (CTGV) developed curriculum around scientists' real world problems. They found, in comparison studies, that students who learned science with tasks that were similar to scientists' actual practices exhibited better conceptual understanding about the content addressed in the tasks, as well as an increased interest in learning about science (Edelson, Gordin, & Pea, 1999). Linking students with actual scientists through networked communication for the purpose of data sharing has been shown to lead students into discussion of more complex content than might be available in textbooks (S. Y. Lee & Songer, 1998).

One inquiry practice in particular, the ability to construct scientific explanations, is often seen as characteristic of deep understanding of scientific content. Simulations of the social construction of knowledge in a professional community, including argumentation and presentation, have resulted in significant gains by students in understanding scientific concepts, using evidence and explaining their ideas logically and uniquely (Bell & Linn, 2000; Scardamalia & Bereiter, 1991). However, this ability to reason scientifically is not a skill achieved overnight. Prior research has indicated some common student difficulties in constructing explanations, particularly students' difficulty in using appropriate evidence (Sandoval & Reiser, 1997). Students have similar difficulty in classroom discussions, where claims are often made with little backing (Jimenez-Alexandre, Rodriguez, & Duschl, 2000).

In order to help students develop this complex practice, previous research has focused on the use of various types of scaffolding, which may help learners complete more advanced activities and engage in more advanced thinking than they might otherwise be able to do on their own (Bransford, Brown, & Cocking, 1999; Wood, Bruner, & Ross, 1976). A number of these studies have centered on written scaffolds provided within the curriculum that support students' construction of scientific explanations (Bell & Davis, 2000; H.-S. Lee, 2003; Sandoval, 2003). Such research has indicated that written prompts, both generic and content specific, can be effective in promoting greater understanding of science content and in producing more coherent explanations.

However, these studies take place within a single curricular unit. Explanation construction is a practice that crosses over all science content areas and is built over time. Can scaffolding designed for a single curricular unit be beneficial over multiple units? This study focused on student responses to scientific questions in three inquiry-oriented curricula within a single year, all of which had scaffolding designed to support students in explanation construction. This study focused specifically on students' responses on hypothesis or prediction questions. Making scientific hypotheses and predictions can be considered a variation of making scientific explanations. Like explanations, hypotheses and predictions must utilize supporting evidence and reasoning to be considered scientifically legitimate. Hypotheses and predictions are similar to explanations except that they occur prior to events or experiments, rather than afterwards.

This study was designed to characterize the nature of students' hypotheses and predictions across three curricula. The research questions were as follows:

- What are the characteristics of students' predictions and hypotheses across three inquiry-oriented curricula?
- What changes in characteristics do students' hypotheses and predictions undergo between the different curricula?

Theoretical Framework

This study draws on the literature from two primary constructs, scaffolding and scientific explanations, in the design and analysis of this study.

Scaffolding

Previous research has confirmed that the assistance of more knowledgeable other can greatly assist students in conducting and completing learning tasks (Annemarie S. Palincsar & Brown, 1984). This view that learning can be mediated by social interaction is the basis for the theoretical lens of social constructivism. The nature of knowledge from social constructivist perspective is based on five principles: (a) knowledge is not a passive commodity to be transferred from teacher to learner, (b) students cannot and should not be viewed as tabula rasa, (c) knowledge cannot exist separate from the knower, (d) learning in a social process mediated by the learner's environment, and (e) the prior or indigenous knowledge of the learner is of significance in accomplishing the construction of knowledge in a new situation. Social constructivism can trace its roots to the work of Lev Vygotsky. Vygotsky (1978) believed that cognitive skills and patterns of thinking were not primarily determined by innate factors, but rather were products of the activities practiced in social institutions of the culture in which the individual was raised. He proposed that social interaction plays a fundamental role in the development of cognition:

> *"Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological) and then inside the child (intrapsychological). This applies equally to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relationships between individuals."*
>
> *(Vygotsky, 1978)*

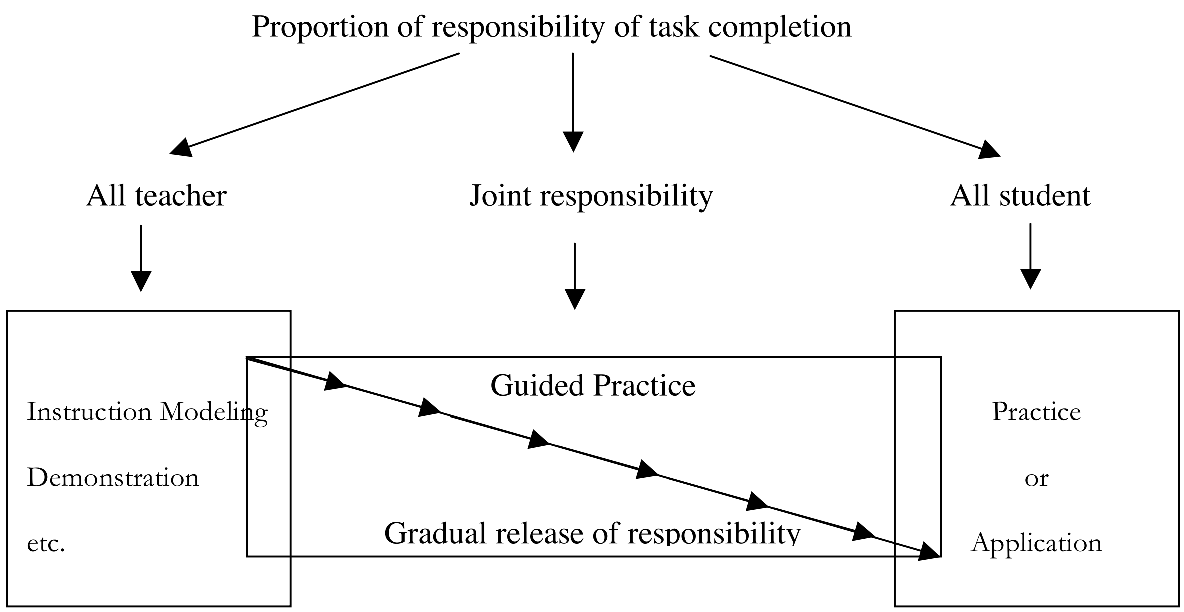
In this view, the assumption is that knowledge is grounded in the relationship between the knower and the known. Knowledge is constructed through social intercourse, and through this interaction we gradually accumulate advances in our levels of knowing. Vygotsky further proposed the zone of proximal development (ZPD) as a visualization of the hypothetical area where learning occurs. Within this zone is the difference between a student's capacity to solve problems on his own, and his capacity to solve them with assistance.

One metaphor for the social interaction that takes place in the ZPD is that of scaffolding. This metaphor was first put forth by Wood, Bruner and Ross (Wood et al., 1976) to describe the social context surrounding learning with assistance. They describe a learner as gaining much more from the other than simple imitation. The scaffolding process:

> *"enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts. This*

> *scaffolding consists essentially of a an adult "controlling" those elements of the task that are initially beyond the learner's capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence. The task thus proceeds to a successful conclusion. We assume, however, that the process can potentially achieve much more for the learner than an assisted completion of the task. It may result, eventually, in development of task competence by the learner at a pace that would far outstrip his unassisted efforts.*
>
> *(Wood, Bruner, & Ross, 1976)*

An important principle of the scaffolding metaphor is that appropriate levels of support be given to the learner and gradually withdrawn as the learner gains knowledge and experience. Four key features can characterize scaffolding (Stone, 1998). The first is the learning context. Initial establishment of the learning context for the interaction between the student and the "tutor" is vital. This both situates and clarifies the learning task for the student and the tutor. The second feature is adequate administration. In order to provide appropriate levels of support, it is necessary to carefully monitor the learner's progress. The third feature is variability. A variety of different supports should be available to the learner. And finally, the fourth feature of scaffolding is withdrawal, sometimes referred to as fading. The support provided by the more knowledgeable other must gradually be faded away if transfer of responsibility from the tutor to the student is to take place. This gradual withdrawal of support can be visualized in the following model below.



(Cazden, 1988)

Scaffolding Student Learning

Scaffolding can be delivered in a number of methods. Generally, we think of a person as taking the role of the more knowledgeable other, e.g. a teacher, a parent, a capable peer. This type of scaffolding would be provided through verbal interaction between the tutor and learner. The tutor, e.g. the teacher, would continuously observe the student and provide the appropriate support based on her observations. This is the model for the enhancement of student reading comprehension through use of the reciprocal teaching method (Palincsar & Brown, 1984). In studying student-teacher interactions, Palincsar and Brown (1984) found teachers initially meet students' needs by providing modeling of appropriate practices, opportunities for practice, and feedback. As students became better at the task over the course of several lessons, the teacher would then scale back her role until the student had taken on the expert responsibilities and the teacher was merely a supportive audience. This study, though widely considered quite successful, is enormously labor intensive, and may be difficult to scale up to larger numbers of students. Ideally, teachers modify and fade scaffolds based on the needs of individual students. When only one teacher is available for an entire class of students, there is concern that teacher-student scaffolding cannot be effective (Stone, 1998).

Another type of scaffolding to consider is the support for student learning that can be provided by written or electronic materials. Certainly curriculum-student interactions are not dynamic or sensitive to student needs in the ways of teacher-student interactions. Nevertheless, they can still be considered a limited use of the scaffolding metaphor (Stone, 1998). Some researchers have argued that limiting the conception of scaffolding to only interactions between individuals is an artificial constraint, and to recall that "ZPDs include not only people but also artifacts…" (Annemarie S. Palincsar, 1998).

The work of several researchers has indicated that other resources can serve some function as more knowledgeable others. Written curricula using reflective self-assessment at the end of each learning task has been shown to help students reflect on the inquiry process involving the modeling of Newtonian mechanics (White & Fredrickson, 1998). Scaffolding can also be embedded in technological resources. In some cases, students are given a range of supports from which they can choose what they need (Bell & Linn, 2000; Davis, 2000). In other cases, technologies guide students based on their responses to set tasks (Scardamalia & Bereiter, 1991). This study examines the possible effect such written scaffolds within a curriculum might have on students in their development of scientific explanations, hypotheses, and predictions.

Scientific Explanations

The formulation of scientific explanations serves important learning functions in the science classroom. Students' creation of external representations, such as explanations, makes ideas perceptually salient, focuses inquiry activity, and teaches process skills (Toth, Suthers, & Lesgold, 2002). Explaining "enables one to reason more logically and scientifically, promotes understanding scientific theories within domains, fosters understanding about why problems are formulated as they are, and, most important, clarifies what needs to be explained" (Coleman, 1998) Similarly, the clarity and coherence of student explanations can serve as an assessment of student understanding (Metz, 1991; Woodruff & Meyer, 1997). As such, developing students' ability to make explanations should be a fundamental goal of science education (Dagher & Crossman, 1992; Edgington, 1997; Sandoval & Reiser, 1997; Woodruff & Meyer, 1997).

The importance of scientific explanations is also reflected in national reform documents regarding science education (American Association for the Advancement of Science, 1993; National Research Council, 1996). The National Research Council (2000) incorporates explanations into four of their five essential features of classroom inquiry.

Though the value of scientific explanations in the classroom and even in scientific practice is not in question, the precise definition of a scientific explanation is not clearly established. There are, however, common elements of agreement, regardless of where the definition arises. First, and most obviously, is that an explanation must go beyond mere description (Woodruff & Meyer, 1997). In description, qualitative information is presented. In explanations, a mechanism must be proposed, and connections drawn between the information presented. Another common feature of explanations is that causal relationships must be utilized (Sandoval & Reiser, 1997). Explanations explicitly link events together in logical cause-and-effect relationships. Additionally, those causal mechanisms must be substantiated by observed data (Kuhn, 1989; Sandoval, 2003). Explanations must cite relevant data, and make valid inferences from that data, that supports the claims being made.

In building a course of reasoning for the purpose of persuading an audience to your point of view, scientific explanations can be considered similar to rhetorical arguments. Toulmin (1958) presents a model for rhetorical argument, which contains the following elements:

Claim:	a conclusion statement whose truth is to scrutinized in the argument
Data:	facts that are used to support the claim
Warrants:	reasons proposed to justify why data guarantee the claim
Backing:	the source that authorizes the validity of warrants
Qualifiers:	the strength of the claim conferred by the warrant
Rebuttals:	conditions where the warrant is not held true

This model is adapted by many science educators in developing a framework for scientific explanation (Bell & Linn, 2000; Driver, Newton, & Osborne, 2000; Sandoval, 2003). Hypotheses and predications can fall within the same framework. They share many of the same traits as explanations (claims, supporting statements, and reasoning for those claims) with the exception that the claim is speculative rather than conclusive. The framework adopted by this study includes two components for hypotheses and predictions, namely (1) the coherent articulation of causal claims, and (2) the use of evidence to support those claims (a combination of Toulmin's data and warrants).

Even with this simpler model, the formulation of explanations logically and consistently using relevant evidence is a difficult task for many students (Bransford et al., 1999; Butcher & Kintsch, 2001; Kuhn, 1989). To begin with, teachers and/or textbooks often provide extant explanations (Kuhn, 1993). As a result, students are not typically given much opportunity to practice building their own explanations (Chi, Leeuw, Chiu, & Lavancher, 1994). Even when students do engage in their own explanation building, studies in middle school have found that students' natural inclination is to provide more detailed descriptions, or engage in teleological reasoning, rather than provide causal links in their explanations (Wong, 1996; Zuzovsky & Tamir, 1999). These same studies find that student explanations are often incomplete, imprecise, and contain many implicit assumption, rather than explicit conditions.

High school students sometimes fare little better, though older children do appear to identify more sophisticated relationships that those of younger children. Sandoval and Reiser (1998, 2002) have found that even when students do appear to use evidence to construct their argument in their own minds, they do not use that evidence to substantiate claims when making their explanations to others. Sandoval and Reiser speculate that this may be because students feel that the relationship is obvious.

On the whole, many of those who have studied student explanations have hypothesized that lack of experience with explanation as a scientific discourse activity is one of the underlying reasons for students' difficulties (Hawkins & Pea, 1987; Sandoval, 2003; Unsworth, 2001; Wong, 1996). The discourse of science is highly specialized and sometimes even at odds with students' native discourse (Jegede & Aikenhead, 1999). Students will not bridge this gap without guidance, teachers can help students navigate through changes in discourse practice (Lemke, 1990). Enculturating students into using the explicit, causal discourse that is typical of science often requires explicit instruction.

Instructional strategies typically used to support student explanations in science have fall into three main categories: (1) discourse-based strategies/direct teaching used by teachers, (2) conversational prompts used in context of student group work, and (3) content and rhetorical prompts used in the context of written explanations, sometimes embedded within software tools. The first of these strategies has been described at length, though there is little evidence as to their effect. Some teachers use the analysis of explanations provided by the textbook as a way foster student understanding of explanations as a form (Unsworth, 2001). However, teachers' own scientifically inappropriate use of explanations may sometimes confound students, as when teachers use the terms "explain" and "describe" interchangeably (Dagher & Crossman, 1992; Horwood, 1988). On the other hand, teachers who do understand the distinctive characteristics of scientific explanations can engage in discourse moves to support explanation construction. Such discourse strategies include general elaboration prompts (e.g. "tell me more?" or "why?"), specific elaboration prompts (e.g. "so you think it's because of *x*, so does that connect to *y* in any way?"), restating the driving question so that students are given a chance to reflect on whether their explanation fully address the issue, and synthesizing and revoicing student remarks, especially in scientific causal term thus modeling for students to appropriate way to state their own reasoning (Tabak & Reiser, 1999). Finally, some teachers will explicitly discuss the criteria for a good explanation with the class as part of science instruction (King, 1994; Sandoval, 2003).

The development of deeper understanding of content through explanation activities among small groups has resulted in the study of conversational prompts for students. These prompts, given on cards, give students a set of questions to ask each other when working in small groups. The question prompts tend to be similar to discourse strategies used by teachers described above, for example, "How are x and y similar?" or "Can you explain that using the information we learned in class?" and so forth (Bereiter & Scardamalia, 1987; Coleman, 1998; King, 1994).

In written explanations, typically considered an end product in classrooms, multiple prompts can be presented to students simultaneously. Bereiter and Scardamalia (Bereiter & Scardamalia, 1987) describe student explanations as occupying two hypothetical spaces, the rhetorical space and the content space. In the rhetorical space, the coherence and clarity of the writing in its function as a communication with the reader is of primary concern. In the content space, the relevance and validity of the information/data presented is the focus. Both these spaces are vital and must interact closely for a good explanation to be constructed. Butcher & Kintsch (2001) used both rhetorical and content prompts with undergraduate students and found that, "[u]se of content prompts results in clear and immediate benefits to time spent in the writing process stages and to the quality of the text that is produced" (p.317). Software tools such as Explanation Constructor (BeGUILE) (Sandoval & Reiser, 1997), Mildred (KIE) (Davis, 2000) and SenseMaker (WISE) (Linn, 2000) have also been used to provide hints and prompts as to what information to evaluate and include explanations, as well as provide students with a visual space in which to organize the structure of their explanation.

In the units used in this study, similar written scaffolds were provided in the each of the three curricula. These scaffolds included content prompts, sentence starters, and organized space to construct responses. Each curriculum provided students with five to ten opportunities to make scaffolded hypotheses or predictions. There were additional opportunities for scaffolded explanations as well. Unit assessment items included limited sentence starter scaffolding, but no content prompts.

Context

This study took place in a large urban school district in the Midwest. Participants included 3 teachers and 247 students in three public middle schools. The majority of these students were African American or Hispanic from low-income families. The teachers enacted three inquiry-oriented curricula in their grade six classrooms. The first, *BioKIDS: Kids' Inquiry of Diverse Species* (BioKIDS) (Songer, Huber, Adams, Chang, Lee, & Jones, 2002), is an animal biodiversity curriculum that is designed to run for eight weeks in the fall. In this curriculum, students use palm-based software to collect population and habitat data on animals living in their schoolyard. Using this data, students ask questions and investigate animal diversity, interaction, and habitat needs. *Kids as Global Scientists: Weather* (KGS) (Songer, Devaul, Hester, Crouch, Kam, Lee, Lee, & Vekiri, 2001) is a technology-rich weather curriculum that is designed to run eight weeks in the winter. In this curriculum, students use a CD-ROM to access actual weather data from around the country. Students investigate concepts of temperature, air pressure, humidity, and weather fronts in activities that use data from weather systems around the country. *How Do Machines Help Me Build Big Things* (Big Things) (Rivet & The Center for Highly Interactive Computing in Education, 2002) is a simple machines curriculum designed to run for eight weeks in the spring. In this curriculum, students visit or discuss a local construction site, conduct mechanical advantage experiments with a variety of simple machines, and culminate their experience by designing and building a complex machine that could be used in building construction. All three curricula were designed to foster student inquiry by focusing on three areas of inquiry skills:

1. The formation of scientific explanations from evidence
2. The analysis of various types of scientific data (charts, graphs, maps)
3. The building of hypotheses and predictions (based on relevant evidence)

Explanations, hypotheses, and predictions were scaffolded in a similar manner across all three curricula, with sentence starters and content prompts as shown in the figure below.

Sometimes biologists get phone calls or emails from people asking them to identify an animal. How would you respond if you got the following description?

Harry found a small creature in his yard. As shown below, it has a hard body and lots of legs – he counted 14 of them. Harry thought it's a pillbug, and pillbugs are insects.

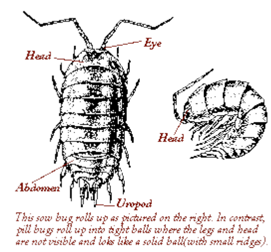


Diagram is from http://insected.arizona.edu/isoinfo.htm

1. In the spaces below, fill in your claim and explanation. Use **Hints.** if needed.

Your Claim: **I think Pillbugs** **<u>are | are not</u>** **insects**
(Choose one)

Explanation: **because…**____

____

____

Hints. In your explanation:
- Think about how many body segments the pillbug has.
- Think about how many legs the pillbug has.
- Think about what the pillbug's body covering is like.

Data Collection

This study analyzed the hypothesis or prediction responses on the pre and post tests of 135 students throughout their entire sixth grade year. Additionally, 12 students were selected for in-depth interviews. Six students were male, and six were female. These students were spread across three different teachers at three different schools within the district. The respective teachers identified three of the students as high performing, six as average achievers, and three as relatively low performing students, based on classroom achievement and performance. All students participated in all three of the curricula: BioKIDS, KGS, and Big Things. BioKIDS was enacted in the fall term, from September through November. Students then engaged in a traditional curriculum on light and sound. KGS was enacted from February to mid-April, and Big Things was enacted immediately following from mid-April until the end of the school year in June. Though each curriculum focused on different science content, all the programs were designed to foster scientific thinking and reasoning. In each curriculum, students collected and analyzed data related to the content area. Scaffolds were provided to help students formulate explanations of scientific questions and concepts using their observations and measurements as evidence.

One primary data source for this study was selected student responses from the pre- and post-tests administered for each of the three curricula. A total of fifteen items, five items from each assessment, were selected for analysis. The focus of these assessment items was for students to demonstrate the ability to make hypotheses or predictions that are justified by evidence either provided by their assessment item or from their own content knowledge background. The fifteen items ranged in complexity. Six items were multiple-choice items, where students were asked to select the most scientifically appropriate hypothesis or prediction from the choices provided and nine items were open-ended, where students were expected to make their own hypothesis or prediction and provide supporting evidence and reasoning.

The other primary data source was student interviews. Interviews were administered at the conclusion of the simple machines enactment. During the interview, students were asked to respond to twelve items from the three curricular assessments using a think-aloud procedure. The purpose of these interviews was twofold. First, to further probe student reasoning on test item responses. Second, to obtain reasoning from students for the several test items that were multiple-choice. Each interview took approximately 30 minutes. All student interviews were digitally recorded and transcribed. Excerpts from interviews are used in this paper to support interpretation of the assessment item analysis. Student notebook responses were also collected for all work done during the enactment of the three curricula, and were used as a secondary data source.

Methods

Assessment items, interview responses, and student notebooks were analyzed using the same set of codes and coding rubric. To identify each response and identify patterns of responses in the students, a first pass set of codes was determined for each interview item response using a grounded theory approach (Strauss & Corbin, 1990). After reviewing dozens of responses, two general patterns: scientific accuracy of hypothesis/prediction, and internal consistency between hypothesis/prediction and the supporting evidence and reasoning. Each response was coded at two levels. First, a code was used to identify responses supported by evidence (table 2). Once responses were identified, a second set of codes was used to characterize the nature of the supporting evidence, if provided (table 3). Overall, for each response, the presence of a claim, the presence of supporting evidence, the type of evidence utilized, and overall accuracy of the hypothesis or prediction, were noted.

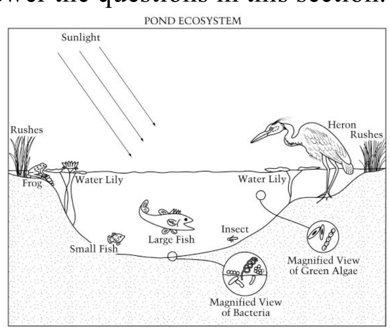
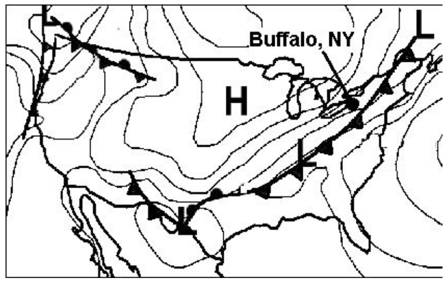
Code	Description	Sample assessment item	Sample response
4	Internally consistent	The picture below shows a pond ecosystem. Use this picture and what you know about the things in it to answer the questions in this section.  If all of the small fish in the pond system died one year from a disease that killed only the small fish, what would happen to the algae in the pond? Explain	The algae in the pond get more and more. The little fish used eat the algae and since they dead, they nothing to stop the algae from growing
3	Internally inconsistent		The algae will all die. When the little fish are gone, they'll be nothing to eat the algae, so the algae will be gone.
2	Unsupported correct		The algae will get big and multiple
1	Unsupported incorrect		The algae will all be gone
0	No response		

Table 2. Coding key used to identify evidence-supported responses

Code	Description	Sample assessment item	Sample response
1	Scientifically plausible	Omar and Norma are planning to go on a picnic today. They look out of the window and see some high, thin clouds. Choose an answer and complete the sentence below to explain your answer. I think it <u>will rain | will not rain</u> because…	I think that it will not rain because the clouds Omar sees are thin, high clouds. If it's going to rain, the clouds be thick and dark. These clouds are not rain clouds.
2	Scientifically implausible		I think it will rain because high, thin clouds bring rain
3	Appropriate evidence	The pressure map below was constructed on March 2nd, 2003. On this date,	I think the air pressure will be high because the

<table>
<tr><td></td><td>used</td><td rowspan="3">Buffalo, NY, had heavy snow with overcast skies. The temperature was 0 °C, and the pressure was 1008 mb at 1 PM.


Based on the map above, predict pressure in Buffalo, NY, on March 3rd at 1:00 PM. Give one reason that supports your prediction.</td><td>H means high pressure and it moving toward the city</td></tr>
<tr><td>4</td><td>Inappropriate evidence used</td><td>The air pressure will increase because the map shows a cold front coming toward Detroit</td></tr>
<tr><td>5</td><td>Extraneous evidence used</td><td>The air pressure will go down because there’s an ice storm coming, and when the weather is worse and that means the pressure dropped</td></tr>
</table>

Table 3. Coding key for characterization of evidence used to support responses

Results

Changes in use of evidence

Over the course of the curriculum, the data show an overall, and quite noticeable, increase in the use of evidence to support hypothesis responses. Figure 1 shows the percentage of assessment responses where students provided evidence in support of their hypothesis or prediction. This does not include multiple choice questions, where no opportunity to provide evidence was allowed. This also does not represent whether the claim or evidence was scientifically appropriate, only if students utilized evidence of any kind to support their hypothesis or prediction.

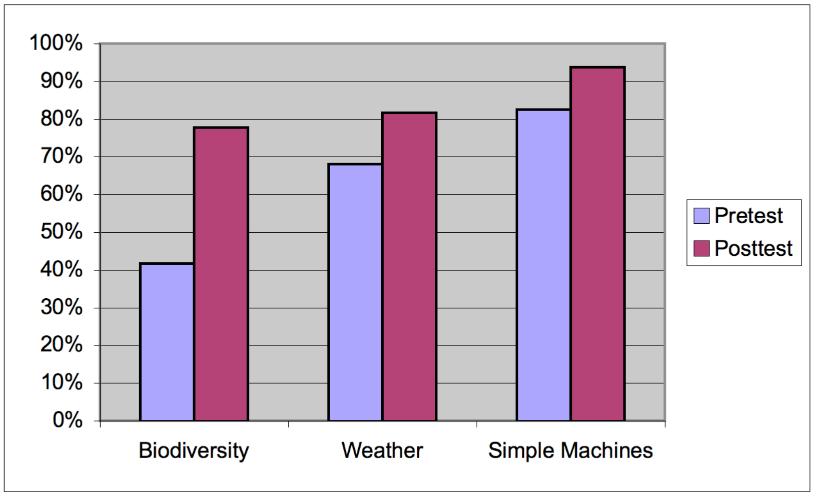


Figure 1. Percentage of evidence-supported responses in three unit tests

The data show that in each of the curricular assessments student use of evidence in support of their claims increased from pre-test to post-test. In addition, utilization of evidence generally increased across the three curricula as well. In each of the pre-tests, evidence use has consistently increased, from 41.7% in the biodiversity curriculum, to 68.0% and 82.4% in the weather and simple machine curricula.

In post-enactment interviews, 10 of 12 students made specific mention of the need for evidence when answering open-ended questions, as demonstrated in excerpt below:

> I: Let's look at question number six then. The pressure map below was constructed on March 2nd 2003. On this day Buffalo, New York had heavy snow with overcast skies. The temperature was 0ºC and the pressure was 1008º, I'm sorry, 1008 millibars at 1pm. Based on the map predict the pressure and cloud condition in Buffalo, New York on March 3rd, the next day at 1pm. What do you think the pressure might be like tomorrow, the next day?
> S: Um, I think it'll be 100 and 10
> I: What?
> S: 1010, I think maybe. Because there's a lot of cold front coming up but then it'll probably will go up. I'm not sure if they go together like that.
> I: You're not sure about…?
> S: Do the cold fronts come with more pressure. It might be the other way.
> I: So, would you put that down? I mean, what would you write as you answer then?
> S: Yeah, I say it would go up because of the cold front coming.
> I: But you're not sure about the cold front and higher pressure going together?

S: Yeah, it might the opposite. I forget, but you still got to say the why. Like in a claim-evidence, you has to have the because part.

Here the student references identifying this question as being similar to the scaffolded questions in the curriculum (referred to by teachers as "claim-evidence" questions), necessitating the need to provide some evidence.

Content vs. Consistency

Figure 2 shows the percentage gains in student use of scientifically appropriate evidence in making hypotheses and predictions in the three assessments. In scientifically appropriate responses, students have made a scientifically accurate hypothesis or prediction, and supported with factually accurate evidence.

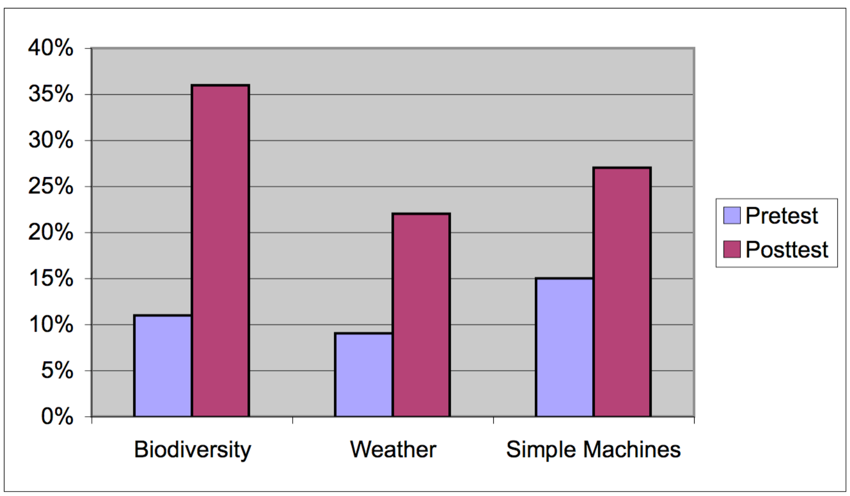


Figure 2. Percentage of scientifically correct responses across three unit tests

In figure 2, we see the greatest gains between pre- and post-test in the first curriculum. Students increased their percentage responses with identification of appropriate evidence from 11% to 36% in the biodiversity curriculum, a 24 point improvement. In the weather curriculum, students gained only 13 points over the initial 9%, and in the simple machines curriculum, there was only an 11 point increase. Similarly in figure 3, the data show that in the first curriculum, consistent responses increased by 28 percentage points from pre- to post-test. However, in the second and third curriculum, consistent responses increased by only 14 and 8 points respectively.

These differences across curriculum may be attributable to differences in enactment. The curriculum with the greatest gains (biodiversity) also had the most complete enactment, with 98% of the curriculum activities completed. Teachers were only able to partially enact both the weather curriculum and the simple machines curriculum, completing 68% and 66% of curriculum activities respectively.

Figure 3 shows the comparison of consistent responses from pre- and post-test for each of the three assessments. In consistent responses, students provide evidence that supports the hypothesis or prediction claim they make. Unlike scientifically appropriate responses, that evidence may not be factually accurate. However, it does logically support the claim that is being made. For example, if students have the erroneous belief that algae consume small fish in a food chain, they might use that fact as evidence that algae population will decrease if all small fish are killed. This response would be scientifically inaccurate, yet internally consistent.

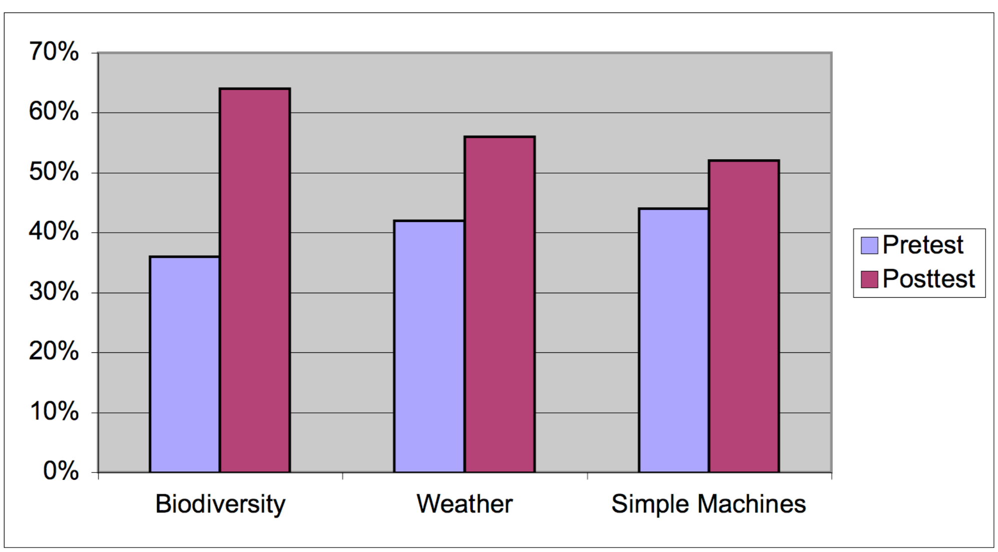


Figure 3. Percentage of consistent responses across three unit tests

When comparing figures 2 and 3, it is clear students are far more likely to produce internally consistent responses than scientifically accurate ones. For example, in the biodiversity unit, almost 64% of responses are consistent, whereas only 36% of the responses are actually scientifically correct. The aggregate of this data, along with the same data for multiple-choice responses is presented in figure 4.

Figure 4 compares students' identification of the scientifically appropriate claim to claims that are supported internally consistent evidence and reasoning, regardless of whether or not they are scientifically accurate hypotheses/predications, this time with addition of multiple-choice items. Student reasoning behind multiple-choice items was established in interviews. These results are averaged across all three assessments.

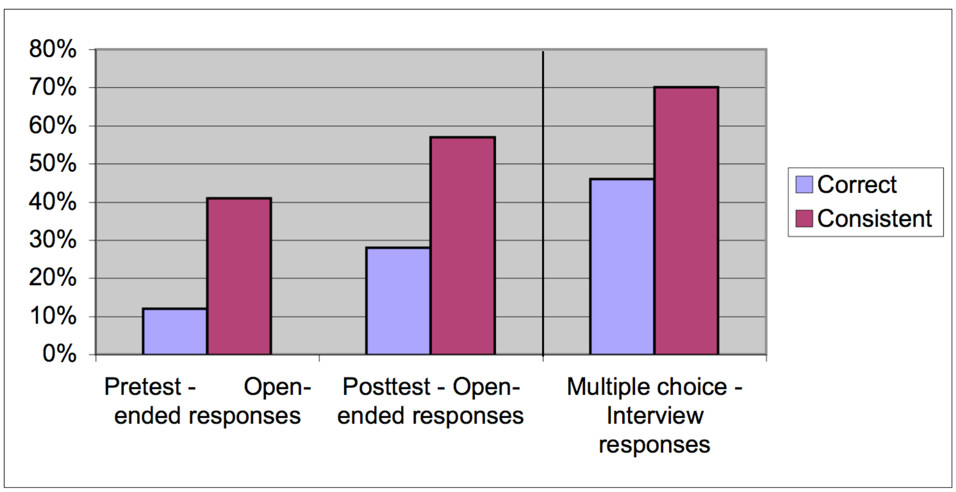


Figure 4. Percentage correct vs. consistent responses across three units

Figure 4 shows that while only 12% and 28% of responses in open-ended pre- and post-test items had scientifically correct claims respectively, 41% and 58% of the claims used internally consistent supporting evidence. Similarly, on the multiple-choice post-test items only 46% of the responses were scientifically correct, but in interviews students were able to provide internally consistent evidence and reasoning for their responses 70% of the time. The following question is an example.

> A meteorologist predicts that the temperature is going to rise in the next few days, and that it will be bright and sunny. Choose the evidence below that can support her prediction.
>
> A. A cold front is moving ahead of a low pressure system
> B. A cold front is moving ahead of a high pressure system
> C. A warm front is moving ahead of a low pressure system
> D. A warm front is moving ahead of a high pressure system

In this question, D is the correct response. The rise in temperature is associated with the warm front and high pressure is associated with pleasant weather. However, many students interviewed selected C as their answer, with the explanation that the warm front would cause the temperature to increase and that low pressure is associated with pleasant weather. In that case, content misconception results in an incorrect response, but reasoning is internally consistent.

Student Trajectories

The differences in evidence-usage, content accuracy, and internal consistency can be seen at the individual level as well. An illustration of such a progression in relatively high-performing student, "Jamar", can be seen in the figure 6 below. Jamar's performance is representative of the three high performing students interviewed in the subsample. Prior to any curricular enactment, Jamar's explanations contained claims, but sometimes without evidence, or directly relevant evidence. For example, below are Jamar's responses to the food web question, "What will happen to the large fish (if all the small fish die from disease)?" in the first curriculum.

> *Pretest: They go hungry.*
> *Posttest: The large fish will starve because the small fishes are died so theyres no food.*

Contrast this to Jamar's pre/post-test responses to a question from the last curriculum. In this question, a brick is shown with an 8N force arrow pulling to the left and a 2N force arrow pulling to the right. Students are asked to predict with way the brick will move.

> *Pretest: I think the brick would not move because it gets pulled on in both directions*
> *Postest: I think the brick would move to the left because there is a greater amount of force on going that way (8).*

Figure 6 shows that over the course of the three curricula, Jamar tended to provide internally consistent responses, even the absence of factually correct responses. As shown above, by the end of the year, the majority of Jamar's responses were internally consistent, both after and before enactment of the final curriculum.

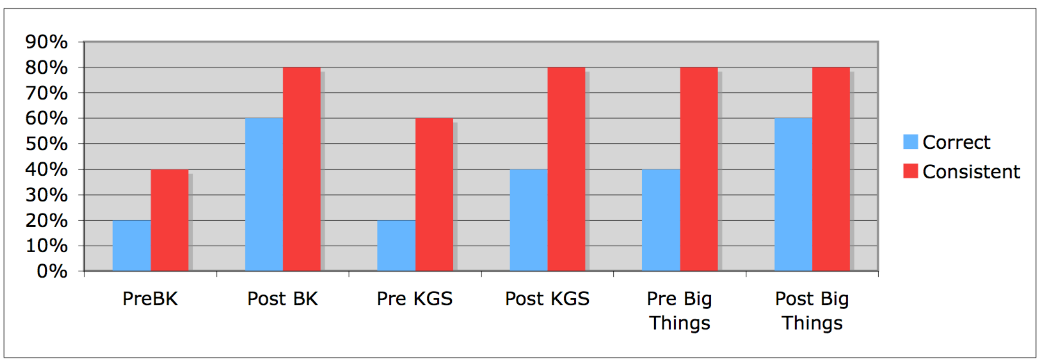


Figure 6: Jamar's hypothesis/prediction response performance across three curricula

Jamar's recognition of the importance of supporting evidence is illustrated here in this excerpt from his post-enactment interview:

> I: Let's take a look this question. In a pond shown here, the small fish eat algae. If the small fish in the pond died one year from a disease that killed only the small fish what would happen to the algae in the pond? Explain why you think so? What do you think?
> S: Well I thinking that the small, the algae ...
> I: Algae?
> S: ... the algae would stay alive or will be eaten by other species.
> I: Okay so those are kind of two opposite things right, one is that it will stay alive and the other is that it will be eaten.

S: Yes.
I: Do you think it is more likely to be one or the other or do you think it's equally likely it will be one or the other?
S: Equal. It's just that I think that because it might stay alive because the small fish doesn't exist anymore and the large fish really doesn't eat the algae. So it would mean that the algae has more time to really live. And, uh, it could take over quick.
I: Okay.
S: But it could be that the water lilies eat the algae.
I: The water lily might eat the algae?
S: Yeah. Or insects. I think maybe they eat the algae, I don't know. But then, the algae, it would be eaten and not stay alive. So, it could might be either one, depending on these, you know, these evidences. If it's the insects and water lilies than the algae will get ate up. But if not, then they'd be okay.

Jamar is not certain of the factually correct information regarding the food web of the pond, but he provides consistent evidence for each of the two possibilities.

Low performing students displayed similar, though much less pronounced gains. Figure 7 below shows the performance of "Tiana", one of the three low-performing students interviewed, over the course of the year. Though Tiana's percentage of consistent responses is a not great deal higher than her correct responses, it is important to note that Tiana also left several of her responses blank, especially in the second two curricula.

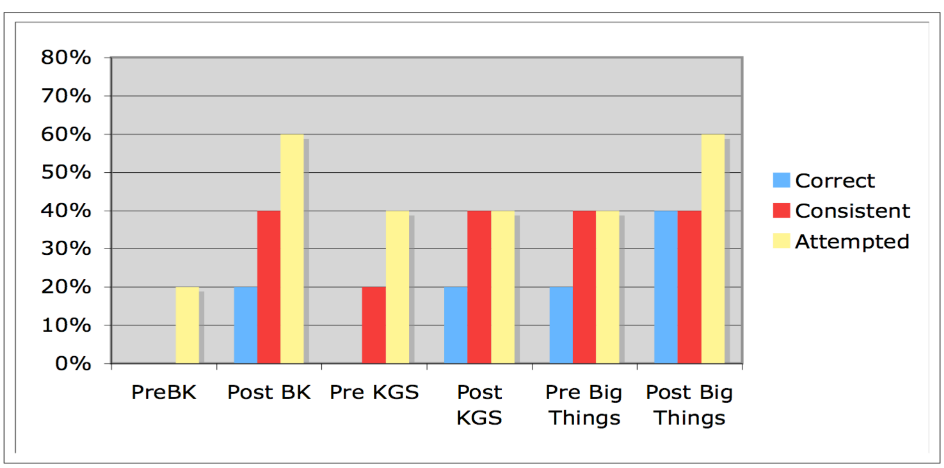


Figure 7: Tiana's hypothesis/prediction response performance across three curricula

Tiana was much less likely to attempt answer questions if she was unsure of the correct answer, which appeared to be often. This can be seen in the following excerpt from her post-enactment interview:

I: Omar and Norma are planning to go on a picnic. They look out the window and see some high thin clouds. What's your prediction about whether or not it will rain?
S: ...
I: Can you make a prediction about whether or not it will rain?
S: No…
I: No?
S: I don't know
I: Well, can you take a guess? What do you think it might be? I mean, if you were answering the question on a test, what would you pick as the answer.
S: I don't know. I'd just leave it 'cause I don't know. Maybe come back.
I: Okay, so if you came back, what would you think?
S: What?
I: About the rain. Do you think it might rain or not rain at the picnic?
S: Oh. I don't… (re-reads question) Well, maybe I think it's going to rain because sometimes you look outside and there's clouds and they say it's going to rain so…
I: Who says it's going to rain?
S: The TV. So maybe it's going to rain 'cause clouds, cloudy and rain.
I: Cloudy means rain?
S: Yeah
I: So, what would you write as the answer to this question?
S: No, I can't write that.
I: No?
S: No, sometimes it's not. Sometimes you got clouds and there's no rain. So, I don't know. I skip that one.
I: Okay.

Like Jamar though, when Tiana did attempt to answer questions, a greater number of responses were consistent compared to those that were factually correct. Tiana has also absorbed the importance of providing evidence from curriculum, illustrated here:

I: Your teacher wants you to figure out a way to increase the biodiversity of your schoolyard. Make a hypothesis of one way to increase the biodiversity of your schoolyard.
S: Oh, oh, you could plant more trees.
I: Okay

S: Where the because? Well, anyway, it's because for more like birds, insects could start living there.
I: Live where?
S: In the trees, there's food and hiding.

Discussion

Two major trends were observed in the data from this study. The first is the improvement in student use of evidence to support claims in their scientific explanations. The trend in figure 1 suggests that over the course of the three curricula students are more accustomed to providing evidence-supported responses. When students do not have the specific content knowledge to make an educated hypothesis or prediction, their responses may vary from wild and irrational speculation to a reasoned guess. Scaffolds in our three curricula are designed to teach students that all scientific claims must be supported by relevant evidence. By the end of the third curriculum, students are providing evidence for nearly all test items, both before and after the curriculum. This seems to indicate that even the absence of content knowledge, students recognize that all scientific claims must be supported by evidence of some kind. The student interviews suggest that the format of scaffolding present in the curriculum may support students' recognition of the need to provide supporting evidence when making scientific claims.

The second major trend observed in the study was the contrast between hypotheses that had scientifically sound claims supported by factually correct evidence and hypotheses that were internally consistent, but not necessarily scientifically accurate. It can been seen in figure 2 that each curriculum resulted in significant student improvement in accurate responses, though clearly students had the greatest improvement in the first curriculum. As mentioned earlier, the improvement from pre to post was not as large in the weather curriculum, but this may be attributable to the fact that the enactment of the weather unit was not as complete as for the biodiversity unit. In some cases, students' lack of content knowledge may mean that they were not aware of the kinds of evidence that could be utilized for making their hypotheses and predictions. This would, of course, negatively impact their ability to make scientifically accurate hypothesis.

Of greater interest is the fact that across the three curricula, more responses are consistent than correct. In the comparison of figures 2 and 3, students were noticeably more able to use relevant evidence in support of their hypothesis or predication than they were to identify a scientifically accurate or plausible hypothesis or predication. That is, student claims tended to be sound, considering the (sometimes incorrect) factual knowledge used as evidentiary support for their hypotheses and predictions. This seems to indicate that student inquiry reasoning skills may be measured separately from content knowledge, at least in part.

Students made gains in consistency over the course of the year, though the initial level of consistency in student responses also began at a higher level. Unlike the use of evidence in responses, student interviews did not indicate any conclusive evidence that the scaffolding in particular helped response consistency. It is not certain what may have contributed to the improvement in response consistency, but the data clearly indicate improvement of the year. This should merit further study.

Overall, the data suggest that students demonstrated reasoning skills even in the absence of content knowledge. It may be that the scaffolding supports students in their reasoning ability, even while it is not sufficient to support their content-specific subject matter knowledge. It does appear that the curriculum scaffolding is effective in supporting student construction of evidence-supported hypotheses and predictions. Subsequent curricula have expanded on the evidence scaffolding, attempting to improve support of appropriate evidence identification.

Conclusion and Implications

Though it is encouraging that students are making internally consistent hypotheses, significant numbers of students are not still not able to identify the appropriate evidence for their claims. This failure may indicate that the content prompts are inadequate. Future curriculum units may wish to feature content prompts more prominently. However, wording of the question or confusion about the relevant concept is just as likely to be the cause of inaccurate evidence identification. For example, a common misidentification occurs in the fishpond ecosystem item shown in table 2. In the item, students are told that the all the small fish in the pond die and are asked to predict what will happen to the pond algae. The desired response is that students predict pond algae will increase because small fish eat algae and in the absence of small fish, the algae will grow unchecked. A common alternative response is that algae will decrease because they eat small fish and in the absence of small fish, they will starve. In this example, students are correctly recognizing that food chain relationships are the relevant concept, but they are mistaken in the direction of the food chain energy flow. It may be the additional activities teaching basic concepts are needed in certain locations, rather than increased explanation scaffolding.

References

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Bell, P., & Davis, E. A. (2000). Designing mildred: Scaffolding students' reflection and argumentation using a cognitive software guide. In B. J. Fishman & S. O'Connor-Divelbiss (Eds.), *International conference for the learning sciences 2000*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with kie. *International Journal of Science Education, 22*(8), 797-817.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience and school*. Washington, D.C.: National Academy Press.

Butcher, K. R., & Kintsch, W. (2001). Support of content and rhetorical processes of writing: Effects on the writing process and the written product. *Cognition and Instruction, 19*(3), 277-322.

Cazden, C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinmann.

Chi, M. T. H., Leeuw, N. D., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Coleman, E. B. (1998). Using explanatory knowledge during collaborative problem solving in science. *Journal of the Learning Sciences: Special Issue: Learning through problem solving, 7*(3-4), 387-427.

Dagher, Z., & Crossman, G. (1992). Verbal explanations given by science teachers: Their nature and implications. *Journal of Research in Science Teaching, 29*(4), 361-374.

Davis, E. A. (2000). Scaffolding students' knowledge integration: Prompts for reflection in kie. *International Journal of Science Education, 22*(8), 819-837.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287-312.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences, 8*(3-4), 391-450.

Edgington, J. R. (1997). *What constitutes a scientific explanation?* Paper presented at the Annual meeting of the National Association for Research in Science Teaching, Oak Brook, IL.

Hawkins, J., & Pea, R. D. (1987). Tools for bridging the cultures of everyday and scientific thinking. *Journal of Research in Science Teaching: Special Issue: Cognitive consequences of technology in science education, 24*(4), 291-307.

Horwood, R. H. (1988). Explanation and description in science teaching. *Science Education, 72*(1), 41-49.

Jegede, O. J., & Aikenhead, G. S. (1999). Transcending cultural borders: Implications for science teaching. *Research in Science and Technological Education, 17*(1), 45-66.

Jimenez-Alexandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "doing the lesson" or "doing science"; argument in high school genetics. *Science Education, 84*, 757-792.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*(4), 674-689.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*(3), 319-337.

Lee, H.-S. (2003). *Scaffolding elementary students' authentic inquiry through a written science curriculum.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI.

Lee, S. Y., & Songer, N. B. (1998). *Characterizing discourse in an electronic community of science learners: A case of the kids as global scientists '97 message board.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, San Diego, CA.

Lemke, J. L. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex Publishing.

Metz, K. E. (1991). Development of explanation: Incremental and fundamental change in children's physics knowledge. *Journal of Research in Science Teaching, 28*(9), 785-797.

National Research Council. (1996). *The national science education standards*. Washington, D.C.: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, D.C.: National Academy Press.

Palincsar, A. S. (1998). Keeping the metaphor of scaffolding fresh - a response to c. Addison stone's "the metaphor of scaffolding: Its utility for the field of learning disabilities." *Journal of Learning Disabilities, 31*(4), 370-373.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoreing activities. *Cognition and Instruction, 1*(2), 117-175.

Rivet, A., & The Center for Highly Interactive Computing in Education. (2002). *How do machines help me build big things?* Ann Arbor, MI: University of Michigan School of Education.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences, 12*(1), 5-51.

Sandoval, W. A., & Reiser, B. J. (1997). *Evolving explanations in high school biology.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge of the design of new knowledge media. *Journal of the Learning Sciences, 1*(1), 37-68.

Songer, N. B., Devaul, H., Hester, P., Crouch, S., Kam, R., Lee, H.-s., et al. (2001). *Kids as global scientists: Weather! An eight-week inquiry curriculum for middle school atmospheric science.*

Songer, N. B., Huber, A. E., Adams, K., Chang, H.-Y., Lee, H.-S., & Jones, T. (2002). Biokids: Kids' inquiry of diverse species, an eight-week inquiry curriculum using simple, powerful, technologies. Ann Arbor, MI.

Stone, C. A. (1998). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31*(4), 344-364.

Tabak, I., & Reiser, B. J. (1999). *Steering the course of dialogue in inquiry-based science classrooms.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). "mapping to know": The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education, 86*(2), 264-286.

Toulmin, S. (1958). *The uses of argument*. New York: Cambridge University Press.

Unsworth, L. (2001). Evaluating the language of different types of explanations in junior high school science texts. *International Journal of Science Education, 23*(6), 585-609.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

White, B. Y., & Fredrickson, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.

Wong, E. D. (1996). Students' scientific explanations and the contexts in which they occur. *Elementary School Journal, 96*(5), 495-509.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology, 17*, 89-100.

Woodruff, E., & Meyer, K. (1997). Explanations from intra- and inter-group discourse: Students building knowledge in the science classroom. *Research in Science Education, 27*(1), 25-39.

Zuzovsky, R., & Tamir, P. (1999). Growth patterns in students' ability to supply scientific explanations: Findings from the third international mathematics and science study in israel. *International Journal of Science Education, 21*(10), 1101-1121.