Running head: UNDERSTANDING SCAFFOLDING

Expanding an Understanding of Scaffolding Theory Using an Inquiry-Fostering Science Program

Hee-Sun Lee

Graduate School of Education

University of California, Berkeley

Nancy Butler Songer

School of Education
The University of Michigan

Acknowledgement:
This material is based in part upon research supported by the National Science Foundation under grant REC-0089283. Any opinions, findings, and conclusion or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. This paper is based on the dissertation study conducted by the first author of this paper. An earlier version of this paper was presented at the Annual Meetings of the American Educational Research Association, San Diego, CA, April 2004. The authors gratefully acknowledge Joseph Krajcik, Betsy Davis, and Philip Myers for their support of this work. The authors also thank Mike Derhammer, Bette Diem, and their students who showed great enthusiasm and provided their perspectives throughout this study.

Abstract

This quasi-experimental study investigated several parameters of scaffolding theory in an 8-week, technology-enhanced, biodiversity curriculum. Two treatments were created to test whether or not written curriculum materials should gradually diminish the amount of scaffolding as students gain experience with explanation tasks in real-world contexts. Forty-eight students in two 5th/6th combined classes were assigned to the two treatments. The consistent support treatment provided three kinds of epistemic explanation scaffolds, e.g. exemplars, questions, and sentence starters, throughout the eleven inquiry situations in the curriculum. In the fading support treatment, these scaffolds were gradually withdrawn over the three curricular phases. Data included pre and post tests, written explanations, and post interview transcripts of selected students. Results demonstrate that both treatment groups exhibited pre to post gains in the knowledge about biodiversity and the ability to match given evidence to a claim. However, as the fading of explanation scaffolds occurred to the fading support treatment, the consistent support group outperformed the fading support group in formulating explanations from authentic data. High ability students appeared to benefit from consistent content support, while low ability students underutilized the explanation scaffolds provided in this study.

Introduction

One of the most challenging goals of the current science education reform is preparing students for lifelong learning (Linn & Muilenburg, 1996; NRC, 1996). In lifelong learning situations that occur outside classroom settings, students need to formulate, conduct, and sustain their own inquiry in real-world contexts (Minstrell & van Zee, 2000; NRC, 2000). Science class provides an ample opportunity for students to acquire and polish independent skills needed for real-world inquiry (Bybee, 2000). To support inquiry reasoning that is both meaningful to students and faithful to the scientific enterprise, the National Science Education Standards advocate the use of authentic science activities that are “similar to those they [students] will encounter in the world outside the classroom, as well as to situations that approximate how scientists do their work” (NRC, 1996, p. 78).

Authentic inquiry is complicated for students to pursue because they often lack much of the domain-expertise required to effectively reason within real-world contexts (Lee & Songer, 2003). Despite epistemological and motivational advantages of using authentic inquiry (Edelson, 1998), the challenge of, for example, determining salient from irrelevant evidence makes it extremely difficult for students to carry out authentic inquiry as is common within the science community (Chinn & Malhotra, 2002; Kuhn, 1989). Research demonstrates that in any given reasoning situation, students may lack content knowledge, inquiry experience, technological resources, professional commitment, or community support (Bransford, Brown, & Cocking, 2000; Edelson, 1998). As a result, students seldom create meaningful inferences from data without adequate support (Keys, 1999). Scaffolding can be one effective instructional means to address these expert-novice differences (Quintana, Reiser, Davis, Krajcik, Fretz, Duncan, Kyza, Edelson, & Soloway, 2004).

The scaffolding metaphor applies to instructional situations where students, following the guidance of the more knowledgeable other, become competent with academic tasks that are initially beyond their ability (Palincsar, 1998; Wood, Bruner, & Ross, 1976). The more knowledgeable other can successfully diagnose the complex needs of students at various stages of the intended learning and employ proper instructional strategies adaptively to their progress (Tabak, 2004). Recent theoretical advances in sociocultural and distributed cognitions allow the expansion of the more knowledgeable other: from human tutors (Wood et al., 1976) to teachers in classrooms (Palincsar & Brown, 1984), peers (Hogan, Nastasi, & Pressley, 2000), technological artifacts (Quintana et al., 2004), and curriculum materials (Cazden, 2001; Davis, 2003; Linn & Hsi, 2000).

Stone (1998) points out that detailed mechanisms of fading in the scaffolding framework are not clearly understood. Bruckman (2000) suggests that answers to when, how, and for whom scaffolding can successfully fade may not be generalizable beyond the initial learning context. Pea (2004) calls for empirical evidence that can test “the claims of scaffolding theory in which scaffolding theory involves specific formulations of what distinctive forms and processes of focusing, channeling, and modeling are integral to the development of expertise” (p. 443). Pea (2004) also suggests such mechanisms should be examined for students with different capabilities on the intended learning.

The study investigated fading mechanisms in scaffolding students’ development of scientific explanations from data they collected as part of an eight-week curricular program, BioKIDS: Kids’ Inquiry of Diverse Species (Songer, et al., 2002). The BioKIDS program, http://www.biokids.umich.edu, is designed to support middle school students’ active pursuit of inquiry problems in biodiversity. Two treatments were designed to test whether or not written curriculum materials should gradually diminish the amount of scaffolding as students gain experience with explanation tasks in real-world contexts. In the consistent support treatment, three types of explanation scaffolds, e.g. exemplars, questions, and sentence starters, were provided throughout the eleven inquiry situations in the BioKIDS curriculum. In the fading support treatment, these scaffolds were gradually withdrawn over the three curricular phases. In comparing these two treatments, this study addressed the two following questions:

- Question 1: How did the learning outcomes of the fading support group compare to those of the consistent support group?
- Question 2: How did students with different levels of expertise relative to content knowledge and explanation generation respond to fading as compared to consistent support?

In the following sections, this paper discusses background literature related to (1) explanation-focused, real-world inquiry, (2) interventions to support students' explanation building process, and (3) scaffolding.

Explanation-Focused, Real-World Inquiry

Situated cognition provides a theoretical framework for the analysis of what constitutes authentic activities in science class. Brown, Collins, and Duguid (1989) state that authentic activities represent "ordinary practices of the culture" (p. 34) where "meanings and purposes are socially constructed through negotiations among present and past members" (p. 34). A focus on selected ordinary practices of the scientific community leads to two possible emphases for authentic science activities for students. McDonald (2004) defines these two emphases as "authentic learning" and "authentic science." In authentic learning, the science activities are situated in everyday life to prioritize students' interests (Keys, 2001; Krajcik et a., 1998; Linn & Muilenburg, 1996; Linn & Songer, 1991). In authentic science, a priority is put on students' experiencing a streamlined version of scientists' research or data-gathering practices (Chinn & Hmelo-Silver, 2002; Edelson, 1998). These two emphases do not always contradict each other

(NRC, 1996). However, it is often difficult to accomplish both goals with any one set of activities.

Though the epistemological and motivational appeal of authentic science learning is understandable (Metz, 2000), using real-world problem contexts in science class presents many challenges. First, real-world science is inherently complicated. Unlike controlled experiments in the laboratory, real-world phenomena involve many variables that can influence the outcomes of investigations. Students need to assess the relative importance of each variable and prioritize major variables to make causal relationships with the observed outcomes. Prioritizing major variables is difficult and requires domain specific expertise (Lee & Songer, 2003).

Second, students lack advanced knowledge and thinking skills necessary to address complicated real-world problems (Edelson, 1998). Differences between experts and novices are well documented including concept understanding (Anazi, 1991; Lewis & Linn, 1994), scientific reasoning (Clement, 1991; Kuhn, 1989; Schunn & Anderson, 1999), problem solving (Chi, Feltovich, & Glaser, 1981), explaining scientific text (Chi, Lewis, Reimann, & Glaser, 1989), and understanding scientific diagrams (Lowe, 1993). Most students have difficulty learning science within real-world contexts without explicit guidance in, for example, recognizing salient data or diagram components (Edelson, 2001; Metz, 2000).

One program that studied the challenges in making 24 hr weather forecasting situations revealed the importance of creating authentic activities that (1) map to real-world contexts that are interesting to students, (2) emphasize problem situations that are more similar to instructional reasoning situations, and (3) provide guidance in determining salient from irrelevant evidence (Lee & Songer, 2003). Bridging cognitive differences between scientists and students is also found in White's (1989) "intermediate abstraction." Taken together, translated real-world inquiry contexts retain the major aspects of scientific inquiry scientists perform, but provide simplified language, easier access to information, and support in determining salient information in order to allow novices to be more successful at complex inquiry (Lee & Songer, 2003).

In scientific inquiry, scientists create, revise, or reject scientific knowledge including theories, models, laws, and explanations (Schwab, 1962). The goal of teaching science through authentic inquiry is not only to polish students' inquiry skills but also to enhance their ability to acquire, apply, and refine their knowledge using scientifically valid methods and techniques. Student artifacts such as models and explanations are good indicators of knowledge development (Spitulnik, Stratford, Krajcik, & Soloway, 1998; Sandoval, 2004). Since a major goal of scientific inquiry is to gain explanatory insight into the physical world (Hempel, 1966; Kuhn, 1970; NRC, 1996), a curricular goal to focus on students' development of explanations presents similar thinking processes inherent in inquiry reasoning as performed by scientists (Sandoval, 2004).

In addition, learning how to explain has other educational benefits. Coleman (1998) suggests that 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" (p. 390-1). However, school science generally does not provide ample explanation opportunities for students (Chi, Leeuw, Chiu, & Lavancher, 1994; Kuhn, 1993). As a result, most students have difficulty formulating logically consistent explanations that connect to their scientific knowledge (Bell & Linn, 2000; Bransford, et al., 2000; Butcher & Kintsch, 2001; Kuhn, 1989) especially with authentic data (Keys, 1999).

Interventions to Support Students' Explanation Building Process

In order to facilitate scientific writing, instruction needs to help students work fluently in both rhetorical and content spaces (Bereiter & Scardamalia, 1987). In the rhetorical space, students need to think how to develop a plausible, convincing argument. In the case of developing scientifically valid arguments, rhetorical support would help students recognize the importance of (1) having a logical consistency between claims and explanations (Bell & Linn, 2000), (2) providing relevant data to justify their claims (Keys, 1999), and (3) communicating ideas clearly to the reader (Scardamalia, Bereiter, Brett, Burtis, Calhoun, & Lea, 1992). Examples of instructional supports designed to work in the rhetorical space may look like, “Make sure you include data to justify your claim,” and “List three pieces of evidence to support your main idea.” Such rhetorical prompts that do not connect to specific content can be called “content-free” or “content-lean.”

In the content space, students need to make domain-specific connections between evidence and their claims (Kuhn, 1989). To assist this aspect of scientific writing, relevant scientific content can be prompted using questions such as “What does hemoglobin transport?” (Chi, et al., 1994) or using sentence starters such as “The normal predator-prey relation is…the factor in the relationship that has changed is…” (Sandoval, 2003). Direct content prompting can be particularly useful in real-world problem solving where students have difficulty focusing on salient features (Lee & Songer, 2003). Directly prompting content can also be an effective strategy for novice students with weak domain expertise (Sandoval, 2004).

In comparing the effects of content prompts and rhetorical prompts on college students’ writing, Butcher and Kintsch (2001) found out 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…for novice science writers, prompting consideration of and decisions about content results in the most powerful and impressive benefits to the writing process and written text quality” (p. 317). Benefits of content prompting relate to students’ increased ability to generate ideas (Klein, 2000; Patterson, 2001). The importance of content support in scientific writing is not surprising as Millar and Driver (1987) point out that:

> [i]t appears that what children notice, what they will do and the interpretations they give depend on the conceptions they use. These in turn depend on their prior knowledge in a particular context or domain of experience. Thus, a pedagogy which focuses primarily on the learning of processes may be fundamentally misguided (p. 51).

This implies that inquiry skills such as explanation and argumentation cannot be properly developed without supporting students' acquisition of domain-specific knowledge (Hodson, 1988; Kuhn, 1992). Even though molecular biologists know what constitutes a good scientific explanation, it is possible that they may not write scientifically valid explanations in the context of high energy physics. Similarly, students cannot write scientifically valid and sophisticated explanations if they do not possess knowledge that can provide a theoretical foundation for their explanations.

Recalling relevant content knowledge is not enough for students to formulate scientifically valid explanations. In fact, students should also be familiar with epistemic forms of scientific inquiry situated in a specific disciplinary context (Collins & Ferguson, 1993). Sandoval (2004) further elaborates the relationship between general epistemological commitments and discipline specific paradigms.

> The criteria that scientists have for what counts as a good theory in their discipline depend upon the questions that they find important to answer, but also conform to more general epistemic criteria, such as coherent causal mechanisms, parsimony, and so on. General `epistemological commitments entail beliefs about what counts as valued and warranted scientific knowledge, and lead to the development of investigative strategies that can produce such knowledge. The canonical strategy of controlling variables across experiments is valued because it allows for the isolation of causal relationships, an epistemic goal. (p.347-8)

Therefore, explanation building can be more effectively supported if scaffolding focuses on both relevant science content and explanation templates to demonstrate how to achieve causal coherence and establish evidentiary support. Simplistic templates, such as proving key terms, may not be enough. For example, Cavallo, McNeely, and Edmund (2003) asked ninth grade students to write an essay by saying that “In your summary, include an explanation of how CHEMICAL REACTIONS may be related with the following terms: atoms compounds chemical change” (p. 589). Cavallo et al. (2003) found that merely giving students key terms without adequate instruction encouraged students to retain misconceptions and misuse the terms.

Even though research concludes that students’ formulation of explanations needs adequate support, it is not clear whether students benefit more from content-specific or from rhetorical support in real-world inquiry contexts. In some studies, explanation-fostering supports focus on either content space or rhetorical space, and their impacts are compared to no support conditions (Chi, et al., 1994; King, 1994; McNeill, et al., 2004). Other studies suggest that to maximize the impact of explanation supports, researchers need to design interventions that address both spaces (Sandoval & Reiser, 1997). Whether students need content-free or –specific prompts may be dependent upon a range of factors including student knowledge, explanation experience, and scientific complexity of the situation that needs to be explained. As for the specificity of scaffolds, Davis (2003) found that middle school students can reflect more productively from generic reflection prompts than from specific reflection prompts that dictate what to reflect.

Scaffolding

In order to differentiate scaffolding from other types of instructional strategies, Stone (1998) identifies four key characteristics. A *learning context* establishes an academic task that is initially unachievable by students. The teacher provides adequate support commensurate to students' progress (*adaptive*). To provide adequate support to learners with varying abilities, the teacher has a repertoire of support strategies and methods *(diverse*). The teacher's support should gradually fade to assure the transfer of responsibility to students (*fading*).

Two distinctive aspects of the scaffolding concept are its emphasis on "fading" (Guzdial, 1994; Pea, 2004; Stone, 1998) and its sensitivity to students' developing knowledge and ability to perform the task (Palincsar & Brown, 1984). Fading refers to the gradual reduction of support by the more knowledgeable agent in successful tutor-tutee (Wood et al., 1976), mother-child (Wertsch & Stone, 1985), teacher-student (Fleer, 1992; Flick, 2000), or expert-apprentice relationships (Brown, et al., 1989). In the design-based research paradigm where curriculum materials or technological artifacts primarily guide student learning, the definition of fading is less distinct. For example, Pea (2004) argues that curricular or technological scaffolds without apparent fading mechanisms may not be considered as scaffolding but as distributed intelligence.

Scaffolding-minded interventions can address *fading* either as a part of the student's learning process or as an explicit, active intervention strategy. One fading approach in distributed learning environments is to leave "fading" entirely to the users of scaffolds such as teachers and students. In many of the current science education research studies focusing on students' development of complex thinking, scaffolds embedded in written curriculum or technological innovations do not fade during the intervention period (for example, Davis, 2003; Quintana et al., 2004; Sandoval, 2003; White & Frederiksen, 1998). Tabak (2004) suggests a synergistic approach to addresses scaffolding at multiple levels, from design to implementation of scaffolds in order to take the ecological complexity of real classrooms into account. Another fading approach is for curricular materials or technological artifacts to fade scaffolds systemically. In this case, fading can be based on the diagnosis of student progress (Anderson, Boyle, Corbett, &

Lewis, 1990) or on predetermined pedagogical decisions that assume the development of expertise (McNeill et al., 2004).

However, clear guidelines for fading mechanisms are not present in the literature and thus need greater specification (Stone, 1998). One explanation for this lack of specificity is the understanding that every child's learning is unique and depends on many instructional factors including leaner characteristics (McNamara, Kintsch, & Songer, 1996), task characteristics (Lee & Songer, 2003), and interactivity with the learner (Bruckman, 2000).

Learner Characteristics

As the teacher's decisions are sensitive to the progress of each individual student, effective fading mechanisms for guiding students in complex reasoning tasks should differ between high and low achieving students (Pea, 2004). This sensitivity may be even more pronounced for elementary and early middle school learners (Resnick & Klopfer, 1989; Kuhn, Black, Jeselman, & Kaplan, 2000). Metz (1997) suggests the importance of higher-order reasoning tasks for younger students,

> [m]ost problematic, the targeting of purportedly elementary science processes for the first years of schools with a postponement of the integrated practice of goal-focused investigations until the higher grades results in decomposition and decontextualization in the teaching and learning of scientific inquiry. As a consequence, young children engage in science activities such as observation and categorization apart from a rich goal structure or overriding purpose, a practice which is detrimental from cognitive, motivational, and epistemological perspectives. (p. 152)

Metz (1995) further argues that higher-order thinking skills can be acquired by students if instruction is designed to address the weaknesses of the students. To support students' independent inquiry, Metz (2000) designed a curriculum that scaffolds domain-specific knowledge, knowledge about empirical inquiry, domain-specific methodologies, data analysis, and tools. Results indicate that even elementary students as young as second grade are able to engage in part of the scientific inquiry process such as designing controlled experiments.

A similar case can be made for low achieving students. White and Frederiksen (1998) implemented reflective assessment into an inquiry cycle model in the ThinkerTools curriculum to foster qualitative understanding of the Newtonian Mechanics concepts. Even though students with all abilities benefited from the ThinkerTools curriculum, White and Frederiksen (1998) noted the benefits for low achieving students. In another case using several instructional modules in the Science, Technology, and Environment in Modern Society (STEMS) project, Zohar and Dori (2003) found that higher-order thinking skills can be taught to students with all abilities and the gap between high and low achieving students can be narrowed with judicious scaffolding.

One of the interesting differences between high- and low-achieving students is in what level of support optimizes their performance. In reading comprehension, McNamara, Kintsch, and Songer (1996) found that high achieving students work best with a version of the scientific text that requires inference-building by the learner, while low achieving students need a version of the text with all inferences provided. A similar result is found in an explanation writing task that uses real-world problems. Rivard and Straw (2000) discovered that high ability students do better when they work independently while low to average ability students benefit from peer discussion. These studies suggest that though students with all abilities can benefit from the same treatment (White & Frederiksen, 1998; Zohar & Dori, 2003), tailoring support to meet the differential needs of students optimizes student achievement (Davis, 2003).

Task Characteristics

To understand scaffolding, Bruckman (2000) notices that "the *content* of the help you receive matters, but the *context* in which that support is situated is also of great importance" (p. 330). The learning context is defined by the characteristics of the academic task given to the student (Doyle, 1983) as well as the type of interactions the student has with the social and physical resources (Bruckman, 2000; Blumenfeld, Mergendoller, & Swarthout, 1987). Task characteristics affect both the content of scaffolds and the duration of the scaffolds. If a learning situation allows students to employ the same strategies during a series of academic tasks, it is likely that students internalize these strategies from repeated uses, allowing the fading of support (Palincsar & Brown, 1984; McNeill et al., 2004). However, real-world learning contexts often do not provide consistent contexts in which students can recognize and internalize strategies (Chinn & Malhotra, 2002). For example, Lee and Songer (2003) studied students' justifications for their 24-hr forecasting in several real-world weather situations. The real-world forecasting task asked students to make predictions on maximum and minimum temperatures, clouds, precipitation, and wind direction. The quality of students' justifications for 24-hr forecasts depended upon how closely the real-world situations mapped onto their content understandings about weather systems rather than the accumulation of forecasting experiences. Therefore, scaffolds that directly reduce task complexity through channeling and focusing are necessary, at least until students possess domain expertise to address the task effectively (Pea, 2004).

Interactivity with the Learner

Distributed cognition increasingly recognizes that scaffolding can occur through various means. First, scaffolding can occur through verbal interaction with the teacher as in the case where teachers observe student's progress and prescribe needed support (Wood et al., 1976). In Palincsar and Brown's (1984) reciprocal teaching, teachers modeled a set of specific comprehension strategies such as summarization with students then very gradually turned over these comprehension strategies to students as competence was realized. One of the challenges in replicating this approach was the necessity of a great deal of teacher attention towards each student, and accurate diagnosis of student progress.

Second, peers can scaffold each other (Bruckman, 2000). Scardamalia et al. (1992) showed the benefits of a student-generated communal database that allows students to share and critique each other's ideas and projects. The student-generated database improved the quality of acquired knowledge, written products, and questions generated for further scrutiny (Scardamalia et al., 1992). The effectiveness of peer scaffolding is explained by the shared communication and learning experience between students (Rivard & Straw, 1999) that is not always present between teacher and student (Bruckman, 2000; Hogan, et al., 2000). However, Blumenfeld, Marx, Soloway, and Krajcik (1996) warn that productive collaborative group work is not a guarantee. "The effects of group work depend on how the group is organized, what the tasks are, who participates, and how the group is held accountable" (Blumenfeld et al., 1996, p. 37).

Third, scaffolding can be delivered through technological resources. In some cases, technologies can include multiple layers of support from which students can choose (Guzdial, 1994; Linn, Clark, & Slotta, 2003). In other cases, technologies guide learners based on the record of their responses to prompts on screen (Anderson, et al., 1990). Sometimes technologies provide fixed scaffolds (Davis & Linn, 2000; Sandoval, 2003). Scaffolding embedded in technological tools can access a large number of students with standardized support. However, even with an exemplary scaffolding design, many scaffolding tools are underutilized by students who lack metacognitive awareness of their own learning progress.

Fourth, scaffolding can be delivered through written curriculum materials (Cazden, 2001). Some curricula are deliberately designed to improve the learner's ability to conduct independent inquiry. For example, White and Frederiksen (1998) use reflective assessment at the end of each inquiry cycle to help students reflect on their inquiry process in the modeling of Newtonian Mechanics concepts. Compared to mentoring, the advantage of scaffolding through written curriculum materials is that many learners can be reached simultaneously to achieve specific learning goals set by curriculum designers. The disadvantage is that written curriculum materials are relatively insensitive to individual variations towards the expected competence.

The Research Context

This section describes the learning context in which two scaffolding conditions were manifested as well as the students and teachers who participated in this study.

The BioKIDS Program

The BioKIDS program (Songer et al., 2002) is a technology-enhanced, inquiry-focused biodiversity curriculum for fifth and sixth grade students. Students are guided through eight weeks of activities that utilize their own data about the biodiversity of their schoolyard towards higher order thinking in science. The topics of biodiversity, food webs, and ecology are the central scientific concepts. The curriculum sequence reflects an inquiry cycle of engage, explore, analyze, and synthesize. In the "engage" phase, students are introduced to the biodiversity concept, data collection methods, and technological resources including CyberTracker (Parr, Jones, & Songer, 2002) and Critter Catalog (Espinosa et al., 2002). CyberTracker is software that enables students to record and organize animal sightings systematically. Critter Catalog is a web resource that provides rich information on local animals including appearance, habitat, food, predator-prey relationships, reproduction, human interaction, and endangerment.

After the "engage" phase, students use their own data to explore and analyze a deeper conceptual understanding of scientific concepts such as biodiversity, including an examination of animal abundance and species richness. Students also evaluate how microhabitats in their schoolyard support the animals. In the "synthesize" phase, each student choose an animal and gather information from Critter Catalog to evaluate the ecological need of their focus animal. Students then determine whether their animals can survive in the schoolyard by comparing what their animals need and what their schoolyard can provide. Students create food webs using the animals they investigated in the previous part to explore concepts such as consumers, producers, decomposers, herbivores, carnivores, and omnivores. In addition, students learn about energy flow, interdependence, and interrelationships among the organisms in their food webs.

Participants

This research was conducted in a K-8 school in the Midwestern United States. Due to the open philosophy in this school, students were familiar with hands-on, interest-based, student-directed projects. Forty-eight students in two 5th/6th mixed classes participated in this study. Mr. Moss taught one class of 28 students, and Ms. Boyle taught the other class of 21 students. Two thirds of the students were Caucasian. The rest consisted of seven multi-racial, five African American, two Hispanic, one Asian, and one American Indian student. Results of a district-wide standardized test indicated that these students performed slightly higher than the state average on the state science test.

Mr. Moss has been teaching in this school for ten years after he acquired a teaching certificate in elementary education. Ms. Boyle has been teaching in this school for more than twenty years and had a teaching certificate in general and elementary education. Both teachers had prior experiences with technology and innovative curricula. Both teachers understood the purpose of this study and expressed their support. All BioKIDS classes were observed by the first author of this paper. Often, the researcher acted as a participant observer helping students conduct investigations and clarifying directions on the worksheets.

Two Treatments

This study was designed to provide empirical evidence on fading mechanisms in a distributed learning environment where written curricular worksheets primarily guided individual students' investigations. Unlike most research on scaffolding that compares student outcomes between with and without scaffolds, this study compared student outcomes and artifacts between with consistent scaffolds and with fading scaffolds.

The academic task of focus was for students to formulate explanations to justify their claims using evidence collected from their inquiry investigations. Three types of explanation scaffolds were identified from the literature. First, questions (Q), similar to those used in Chi et al. (1994) and Butcher and Kintsch (2001), oriented students to focus on a small number of salient features for problem solving. For example, questions provided in classifying animals based on observable characteristics include, "Do they have external skeletons?" and "How many legs do they have?" To compare explanations across a range of biodiversity problems, four salient features were emphasized through questions (Q) in each problem. Second, exemplars (E) gave students an idea of how the salient features mentioned in questions (Q) could be incorporated into an explanation. An exemplar for the classification problem is, "I think a beetle and an ant can be grouped together because they have external skeletons and six legs. These data show that both of them are insects." Third, Sentence starters (S) asked students to fill in their claim and justification as shown in this example, "I think ____ and ____ can be grouped together because *[list relevant data or information]…*" Sentence starters for claims were different across inquiry problems, but were consistent relative to evidence generation, e.g. "because *[list relevant data or information]…*" Questions, exemplars, and sentence starters used in this study are listed in Tables 1, 2, and 3, respectively.

Insert Tables 1, 2, 3, & 4 Here

Using these three types of explanation scaffolds, two treatment conditions were established in terms of how fading of the scaffolds occurred over time. As shown in Table 4, both treatments offered exemplars, questions, and sentence starters in Phase I (Problem 1 to 3).

In the fading support treatment, exemplars were eliminated in Phase II (Problem 4 to 7), and then questions were eliminated in Phase III (Problem 8 to 10). In the consistent support treatment, all three types of scaffolds were provided throughout all phases. This fading mechanism was determined to respond to three assumptions: (1) students need to see and recognize what is expected at the beginning of the task (exemplars); (2) channeling and focusing can reduce the complexity of the real-world task (Pea, 2004; Sandoval, 2003); and (3) scaffolds that address more cognitively challenging aspects of the task need to remain longer (Lee & Songer, 2003).

There are two possible projections on the outcomes of this study. The first possibility is that students' gaining experiences with the explanation building tasks allow content-specific scaffolds to fade. Students with consistent content support on the other hand may consider scaffolds as sources for information throughout the tasks, and therefore do not develop important skills for writing explanations such as prioritizing salient evidence. The second possibility is that a series of complicated real-world inquiry situations presents different challenges to students, which may result in denying students' full engagement with the explanation tasks. Comparable (or higher) achievement of the fading support group to the consistent support group can imply that the intentional fading of scaffolds is beneficial for students' development of independent explanations. If the consistent support group prevails, this research can provide contextual insights towards how and why fading of content scaffolds becomes ineffective.

Data Collection & Analysis

Before the eight-week BioKIDS curricular program began, all students took a two part pretest (content test and claim-evidence test). Students were teamed in groups of four to start the first BioKIDS activity. Each student group was assigned to one of the two treatments based on prior biodiversity knowledge measured on the content pretest and prior explanation ability measured on the claim-evidence pretest. Each treatment had six student groups. There was no statistically significant difference between the two treatment groups on the content test, $t(45) = 1.35$, $p = .25$, or the claim-evidence test, $t(43) = .05$, $p = .70$.

During the BioKIDS curricular program, students engaged in eleven inquiry problems where the two treatments were embedded. In the first BioKIDS activity, students observed animals in their schoolyard. Following this activity, students developed individual explanations without scaffolds to address the pre-treatment problem, “Among animals I saw in the schoolyard, which animals can be grouped together?” Right after the first explanation, students were given a second opportunity to explain the same problem on a new worksheet that featured all three types of scaffolds (Problem 1). From Problem 2 to 10, students received the scaffolds as shown in Table 4. After the treatments ended, all students wrote explanations about the post-treatment problem that did not provide any explanation scaffolds. All students then took the content posttest as well as the claim-evidence posttests. In addition, nine students from each treatment were interviewed. Interviewees represented various levels of knowledge and explanation ability. The sections below describe data sources and analysis procedures.

Content Test

The content test addressed knowledge about the biodiversity concepts in the eleven inquiry problems. This test had 19 multiple-choice and 3 short-answer items taken from released standardized tests such as the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study (TIMSS), and the Michigan Educational Assessment Program (MEAP). All responses were coded as either correct (1 point) or incorrect (0 point). The maximum score for the content test was 22. To demonstrate the pre to post improvement, effect sizes were used. To detect treatment effects, repeated measure MANOVA’s were performed using time (to account for pretest to posttest increase) as a within-subjects variable and treatment type as a between-subjects variable.

Claim-Evidence Test

The claim-evidence test consisted of five items that measured students' ability to use given evidence to justify their claims. Each item asked students to make a claim about a biodiversity problem and explain their claim. Item stems were similar to sentence starters (S) used in the treatment. For instance, Figure 1 shows a question about the adaptation of Viceroy Butterflies. This question presents pictorial as well as textual evidence for students to determine whether Viceroy Butterflies would be eaten by predators. In scoring the claim, one point was given for a scientifically correct claim. In scoring the explanation part, one point was given for each of two correct pieces of evidence. The maximum score for the claim-evidence test was 32. The intercoder reliability was .95. To show pre to post gains and treatment effects, the same statistical techniques were used as described in the previous section.

Insert Figure 1 Here

Explanations

All of the eleven inquiry situations were based on students' investigations in the BioKIDS curriculum. The investigations were completed either individually or in groups, but students were asked to explain their claims individually. Students' explanations were expected to meet the following criteria:

- Explanations need to include evidence relevant to the problem.
- Evidence in the explanation should be scientifically valid to justify the claim.
- A claim can be better justified if a larger number of scientifically valid warrants are presented.

Following these criteria, students' explanations were coded in terms of the number of warrants, the number of valid warrants, and the validity ratio. A warrant was implied in each pair of a claim and a piece of evidence, allowing multiple warrants to exist in an explanation. After a total number of warrants was counted in an explanation, each warrant was examined to determine scientific validity. A warrant was as coded as invalid if:

- (Invalid data) Data are not properly measured or cited.
- (Irrelevant data) Data are irrelevant to the problem.
- (Inconsistent data) The connection between the claim and data is inconsistent.

The number of valid warrants was calculated by counting how many warrants were scientifically valid in each explanation. The maximum number of valid warrants in each explanation was 4. The validity ratio, ranging from 0 to 1, was calculated by dividing the number of valid warrants with the number of warrants. This validity ratio measure emphasized the quality of the explanation as a whole, while the number of valid warrants emphasized the quantity of valid warrants in the explanation. This explanation coding process was described using Mary's explanation:

> I think a roly-poly and a spider can be grouped together because they each have eight legs and they're both invertebrates. They both eat insects. These data show that they are arachnids.

In her explanation, she implicitly made four warrants:

- A roly-poly and a spider can be grouped together because they both have eight legs.
- A roly-poly and a spider can be grouped together because they both are invertebrates.
- A roly-poly and a spider can be grouped together because they both eat insects.
- A roly-poly and a spider can be grouped together because they both are arachnids.

The first warrant is invalid because a roly-poly has more than eight legs. The second warrant is valid. The third warrant is irrelevant to the problem because the inquiry problem asks students to focus on physical characteristics not behaviors. The fourth warrant is invalid because a roly-poly does not belong to arachnids. The number of valid warrants in this explanation is 1, and the validity ratio is .25.

Students' explanations were coded separately by two independent coders. Intercoder reliability in all of the inquiry situations ranged from .89 to .99. To compare explanations of the fading and the consistent treatment groups during the course of this study, repeated measures MANOVA's were conducted on the number of valid warrants and the validity ratio, using time as a within-subjects factor. Time representd a series of pre treatment (without scaffolds), Phase I, Phase II, Phase III, and post treatment (without scaffolds). To create a collective score in each phase, explanation scores over the problems in the same phase were averaged.

Interviews

Interview data provided students' perspectives on how they used explanation scaffolds while they were formulating explanations. During an interview, each student was presented with his/her explanations about the following three problems:

- Problem 2 from Phase I: Are _____ and _______ the same species?
- Problem 7 from Phase II: What kinds of adaptations are used by my animal to survive in its habitat?
- Problem 8 from Phase III: Can my animal live in my schoolyard?

All of the eighteen interviewed students were asked to rank the easiest and the most difficult problems to explain and tell why. In addition, students who received fading support were shown missing scaffolds and asked whether and how they liked to change their previous explanations. Each interview took about 20 minutes. All of the student interviews were recorded on audiotapes and transcribed. Interview segments are used in this paper to support the authors' interpretation of research findings.

Research Question 1: Comparing Two Treatments

This section compares the fading support treatment and the consistent support treatment on the content test, the claim-evidence test, and the explanations.

Content Test & Claim-Evidence Test

The content test assessed whether students acquired knowledge about biodiversity. The claim-evidence test assessed student ability to make scientifically valid links between the given evidence and their claims. As shown in Figure 2, both treatments demonstrated significant pre to post gains on the content test, $F(1, 41) = 14.32$, $p < .001$, as well as on the claim-evidence test, $F(1, 39) = 32.81$, $p < .001$. In addition, the consistent support group demonstrated larger gains than the fading support group on both the content test, $ES = .48$ vs. $ES = .28$, and the claim-evidence test, $ES = .64$ vs. $ES = .44$. There was no interaction effect between treatment and time (pre to post). These results indicate that both groups, to a similar extent, acquired (1) the biodiversity knowledge that was necessary to formulate explanations and (2) the ability to match given evidence to a claim.

Insert Figure 2 Here

Explanations

Figure 3 (a) illustrates the trajectories of the two treatment groups on the number of scientifically valid warrants in explanations. In the pre-treatment problem, the fading support group, without scaffolds, included a larger number of scientifically-valid warrants than the consistent support group. Both treatment groups improved from the pre-treatment explanation to the Phase I explanations. This finding indicates that all three types of content scaffolds were beneficial to students. In Phase II, the removal of exemplars (E) in the fading support treatment was associated with a slight decline in the number of valid warrants. However, the number of valid warrants increased in the consistent support treatment group in the same phase. The decline of the fading support group continued as both questions (Q) and exemplars (E) were withdrawn in Phase III, while the consistent support group continued to improve. In the post-treatment explanation, both treatment groups experienced slight declines from their Phase III explanations as students developed explanations without any types of scaffolds. These trajectory changes over time were statistically significant, $F(4, 144) = 7.29$, $p < .001$. In addition, a statistically significant interaction effect existed between time and treatment, $F(4, 144) = 7.75$, $p < .001$, which reflects the consistent support group's improvement over time versus the fading support group's decline.

Figure 3 (b) shows the trajectories of the two treatment groups on the validity ratio. The validity ratio of both treatment groups improved over time, $F(4, 144) = 12.51$, $p < .001$. There was neither treatment difference, $F(1, 36) = .34$, $p = .56$, nor interaction effect, $F(4, 144) = 1.36$, $p = .26$. As students' experiences with explanation building tasks increased, both treatment groups appeared to evaluate each piece of evidence more carefully to determine whether it could support their claim. However, it is noticeable that while the consistent support group continued to improve the validity ratio during the entire treatment period, the fading support group did not improve after Phase II. This finding may suggest that questions (Q) were essential for students to make scientifically valid connections. Without questions (Q), students might have worked harder to select what data to emphasize and how to make connections between the data and their claims.

Insert Figure 3 Here

Discussion

Students, regardless of treatment type, demonstrated pre to post gains in knowledge about biodiversity as well as the ability to match the given evidence to their claims. However, the consistent support group came to make a larger number of scientifically valid warrants than the fading support group. The improvement of the consistent support group and the decline of the fading support group coincided with the gradual withdrawal of scaffolds in the fading support treatment. Student interview data provide insights to understand this outcome.

Responding to diverse needs of students. Two important skills students need to formulate explanations in real-world inquiry contexts are (1) focusing on salient evidence of the problem (Lee & Songer, 2003; Sandoval, 2003, 2004) and (2) building consistent, logical arguments between data and a claim (Bell & Linn, 2000; Keys, 2001; Sandoval, 2003). Each student's progress towards perfecting these abilities is potentially unique. Some students may need content support that identifies salient features all the time, while others can formulate explanations on their own after only a few explanation building practices. In this study, consistent content support better responded to the diverse needs of students who could be in different stages of acquiring explanation skills than the fading content support. Students with consistent support appeared to spend more time thinking about connections between particular data and their claims, instead of examining data from various sources and determining which to use in their explanations. In addition, the predetermined fading order might not adequately reflect students' progress towards independent explanation building as well as their preference for the type of the content support they needed. The short treatment duration might also contribute to the unsuccessful fading of content support.

Even though scaffolds were presented on the worksheets, whether and how individuals used them depended on students' own assessment of what they needed. This study suggests that some students ignored scaffolds on their worksheets if they were confident about their answers for the problem. Ted in the consistent support group said that he did not consider exemplars (E) or questions (Q) because he knew the answer immediately after reading the problem:

I: Did you read the example?
Ted: No, I don't think I read the example…

I: Why this time you didn’t read it? [he read the exemplar provided in the previous problem]
Ted: Well,…because I like knew what I was gonna write and I just wanted to write about it…and I didn’t wanna read the example to guide me.
I: So, you didn’t want to be influenced by…
Ted: Ya, influenced.
I: Did you read the questions under data or evidence?
Ted: No, with the basically same reasons…because I already knew what I was going to write and I knew it would be a good explanation.

However, this high level of confidence was not always correlated with a high quality explanation. Ted was a student with low prior knowledge and explanation ability who indicated that his animal, the Western Chorus Frog, could not live in his schoolyard because “there is no body of small water and no wooded or herb area in the schoolyard.”

In contrast, Carla, a student with high prior knowledge and explanation ability, said that even though the answer was obvious in some cases she used all of the scaffolds and formulated strong explanations such as the following:

> Carla: I think my animal (Walleye) cannot live in my schoolyard because my animal needs deep lakes to live in. There is not a consistent supply of water in our schoolyard. They also eat animals that live in the water and they need food to live. There are no places for a walleye to hide since there is not much water. Also to hide from predators they use weeds, there aren’t any water weeds in our schoolyard. So that’s why they cannot live in the schoolyard.

As most students at this age are known to have difficulty assessing their own learning without explicit support (Davis, 2003; White & Frederiksen, 1998), students in this study might not accurately determine what kind of support they really needed. This finding indicates that when scaffolds are fixed features of the distributed learning environment, how to improve students’ metacognition should also be considered (Pea, 2004).

Alleviating cognitive challenges in real-world problem solving. An important reason why fading support occurs in scaffolding is that what students learned previously can be carried over to the next task (Palincsar & Brown, 1984). If the content knowledge needed for problem solving changes in every explanation situation, it is likely that what students learned through the explanation task in one inquiry situation may not be necessarily translated into the next inquiry situation. Therefore, successful fading of content scaffolds in this study might be too difficult to achieve for the fifth and sixth grade students who developed explanations in the eleven different real-world situations. Each real-world problem addressed different content and/or investigations. Students actually seemed to become more accustomed to connecting a piece of evidence to a claim while they advanced from one phase to the next, as indicated in the improvement in the validity ratio of explanations over time.

However, this study demonstrates that the development of the ability to select a relevant set of multiple salient features in a complicated real-world situation is not straightforward. This interpretation is supported by student explanations about the three Phase III problems. There was a huge treatment difference in the first two Phase III problems (Problem 8 and 9). In Problem 8, the average number of valid warrants in the fading support group was 1.6, while that of valid warrants in the consistent support group was 2.3. In Problem 9, the average number of valid warrants was 1.2 in the fading support group and 2.3 in the consistent support group. However, such difference did not exist in Problem 10 because the fading support group included a similar number of valid warrants to the consistent support group, $M = 2.1$ vs. $M = 2.3$, respectively. Problem 10 asked students to explain possible impacts of a volcanic eruption on the food web in the nearby ecosystem. To solve this problem, there was only one aspect students needed to consider: lack of food. On the contrary, Problems 8 and 9 required students to think about multiple factors such as food, space, and shelter, to determine whether their animals could survive in the schoolyard or how their animals could interact with one another. This finding suggests that students' content support is not necessary when real-world contexts can map onto students' content understanding easily (Lee & Songer, 2003).

Fluent recognition and prioritization of salient evidence are difficult skills to acquire without domain-specific expertise. Many researchers (Blumenfeld et al., 1987; Doyle, 1983, Millar & Driver, 1987) indicate that students can be easily discouraged if what content and how to apply it are not obvious. It was possible that, though students could recollect relevant knowledge about biodiversity, their understandings about the knowledge might not be robust enough to solve some of the real-world problems used in this study.

Interviewees mentioned that both exemplars (E) and questions (Q) were useful. All but two interviewed students thought questions were more helpful than exemplars because questions clearly pinpointed an area of focus. In determining whether the Green Frog and the Northern Leopard Frog are the same species, Helen said that questions helped her because

> Helen: …I might not think about whether they interbreed together because even though it was obvious I wouldn't think about it. And I might not think about, probably about habitat, something that I wouldn't be able to notice or behavior that probably is important.

Students mentioned that exemplars (E) were useful to show how to put what they were thinking into their own words. Exemplars (E) however were not always helpful, particularly when students could not clearly see the connection between the exemplars and their cases. Four interviewed students did not see such connection. For example, the following exemplar was given for the problem of "Can my animal live in my schoolyard?" on the worksheet:

> I think my animal (Gray Wolf) cannot live in my schoolyard because Gray Wolves need to have much larger space than my schoolyard for living and my schoolyard cannot provide enough food, mostly meat, for them.

Matt said that this exemplar was not specific enough for his animal, Earwig, and thus did not help him much:

> Matt: …if you don't find connections with these [examples] and you see that these are totally irrelevant to you because it is [for] Gray Wolf, then it may be hard to find material to write.

Written curriculum materials are not ideal agents to deliver adaptive scaffolds to individual students. Reflecting this view, successful fading of support has been reported in programs that utilized flexible, interactive agents such as teachers (Palincsar & Brown, 1984) or intelligent computer tutors (Anderson et al., 1990). Since each student's need for support differed, the consistent support treatment could meet the needs of all students, regardless of whether all of the content scaffolds were needed. On the other hand, the fading support treatment provided the same reduced scaffolds to all individuals, regardless of individual need or ability. This study suggests that the gradual withdrawal of content support in the written curriculum should be considered only in cases where students have either a flexible and adaptive scaffold system that can adjust to individual difference or when student are considered to have a reasonable amount of experience with highly similar tasks and situations.

Research Question 2: Treatment Difference by Learner Ability

Results discussed in the previous section demonstrate that fading content scaffolds was not an effective strategy in supporting students' formulation of explanations across a range of real-world contexts. This section answers the second research question on how students with different levels of prior knowledge and explanation ability responded to the two treatments.

Three Learner Profiles

Three learner profiles were developed based on students' prior knowledge tested on the content pretest and prior explanation ability tested on the claim-evidence pretest. The mean score of the content pretest, *M* = 16.7, was used to split students into two knowledge groups, high (above average) and low (below average). The claim-evidence pretest scores were divided into three explanation ability groups: high (range = 23-31), medium (range = 19-22), and low (range = 0 - 18). The cross tabulation of these two categories yielded three main learner profiles: 15 students with low knowledge and low explanation ability (L & L), 13 students with high knowledge and medium explanation ability (H & M), and 14 students with high knowledge and high explanation ability (H & H).

Trajectories of Three Learner Profiles

Figure 4 shows the trajectories of the three learner profiles on the number of valid warrants over time. The L & L performances of both treatment groups increased over time in a similar fashion. The number of valid warrants increased from the pre-treatment explanation till Phase II, and did not change much afterwards. As for the H & H profile, the consistent support group included a fewer number of valid warrants than the fading support group in the pre treatment explanation as well as in the Phase I explanations. Then, the number of valid warrants of the H & H profile did not change much in the fading support treatment, while the H & H profile in the consistent support treatment continued to improve. In the post-treatment explanation, the consistent support group outperformed the fading support group. The H & M profile in the consistent support treatment scored initially much lower than the H & M profile in the fading support treatment. The number of valid warrants of the H & M profile increased in the consistent support group, whereas that of valid warrants decreased in the fading support group.

The trajectories of the three profiles on the validity ratio are illustrated in Figure 5. All three learner profiles in the consistent support group exhibited improvement over time. The amount of improvement peaked between the pre treatment explanation and the Phase II explanations and slightly increased afterwards. In the fading support group, the L & L profile and the H & M profile noticeably improved from the pre treatment explanation to the Phase II explanations. After Phase II, the validity ratio of the L & L profile slightly improved while that of the H & M profile slightly worsened. The validity ratio of the H & H profile did not change much over time.

Insert Figures 4 & 5 Here

Discussion

Scaffolding theory (1) assumes developmental trajectories in achieving the final stage of the guided task (Pea, 2004), and (2) expects scaffolds to be employed adaptively to student progress (Guzdial, 1994; Stone, 1998). The breakdown of explanation scores by the three different ability groups shows how the two treatments worked for students who started with different levels of prior knowledge. This section attempts to explain the similarities and differences in the trajectories of the three learner profiles between the two treatments. In discussing each learner profile, student interview data are used to verify the authors' interpretations.

L & L profile. The trajectories of both treatments on the number warrants and the validity ratio were very similar to each other. It appears that students learned to provide at least one piece of scientifically valid evidence over time. However, this lack of difference between the two treatment groups raises a question regarding whether the L & L profile took advantage of the content scaffolds used in this study. Many interviewees with this profile said that some exemplars and questions were confusing. It is possible that students needed to have some understandings of knowledge and a minimum level of explanation ability to make use of the scaffolds. The initial improvement of the L & L profile between the pre treatment explanation and the Phase I explanations could simply be an indication that these students began paying attention to the explanation building task. In fact, most L & L students rarely made scientifically valid warrants in the pre-treatment explanation. As this study began, they recognized the teacher's expectation that they needed to formulate explanations at the end of every investigation. In their interviews, all L & L students, regardless of treatment type, mentioned that Problem 2, the Phase I problem selected for the interview session, was most difficult. See Table 5 (a).

Insert Table 5 Here

Despite the presence of the explanation scaffolds, these students found Problem 2 most difficult, perhaps due to their lack of experience with both explanation building and the content area:

> I: Why do you think the first one was most difficult?
>
> Kate: Because I mean you would have to think about the animals, because you don't know anything about the animal so you have to say, OK, does my animal go with that animal? And then you have to read the questions and examples. Technically, it was kind of hard because…it was like, OK, how…or why do you think that your animal would be good living with that other animal and or the same species, it was kind of hard to put that into words.

As the L & L students repeatedly built explanations from data, they became more comfortable in providing at least one piece of scientifically valid evidence in the problems presented after Phase I. As a result, all interviewed L & L students indicated that Problem 8, the Phase III problem selected for the interview session, was easiest regardless of the existence of the explanation scaffolds.

Results also suggest that the L & L students did not attempt to make as many valid warrants as they could. Instead, they appeared to be easily satisfied with their explanation if they were able to find one piece of evidence to justify their claim. For example, Ted said that Problem 8 was easiest because he was able to find an answer quickly:

I: Why was the last one [Problem 8] easiest?
Ted: Because I already knew the answer from before and I took this and …it was pretty obvious to me that frogs do not live in my schoolyard.
I: Have you not seen any frog in your schoolyard?
Ted: No.

Ted appeared to be more concerned about providing an answer to the problem than thinking deeply to find scientific explanations for his answer. Ted's case is an example of students' lack of epistemic commitments in creating scientific artifacts (Sandoval, 2004).

H & M profile. Table 5 (b) shows the difficulty rating of the interviewed H & M students. Two students in the fading support treatment mentioned that Problem 8, a Phase III problem, was most difficult and Problem 2, a Phase I problem, was easiest. This rating could be explained by their ability to find quick answers:

I: Which one is most difficult?
Holly: on the first one [Problem 2].
I: Why is that?
Holly: ya…because I didn't know that much about the frog and toad… and other two questions are easier because they are on the tiger shark and that was the one I had studied before…so I knew more about it.
I: Which one is the easiest one?
Holly: The schoolyard one [Problem 8].
I: Why this one was easiest?
Holly: Because it was really obvious tiger shark wouldn't live there.

However, their difficulty rating did not agree with the trajectory of the H & M students who received fading support. Figure 4 (a) shows that the number of valid warrants is highest in Phase I and lowest in Phase III. The reduction of content scaffolds is correlated with the decline of the number of valid warrants in the explanations of the H & M students, but not with the difficulty rating. This finding suggests that the difficulty of the problem was rated based on whether they knew the answer rather than whether they could explain. The lack of explanation scaffolds, in fact, was not mentioned as a reason by the two interviewed students.

The difficulty rating of H & M profile in the consistent support treatment was dependent upon their assessment of the problem complexity related to which animal students chose for their investigation and explanation. For example, Earl selected two animals, Gray Wolf and Black Bear, to determine whether these animals were the same species (Problem 2). However, his choice of the Funnel Web Spider made him more difficult to answer its adaptation method in the school habitat (Problem 7). Earl therefore rated Problem 7 as most difficult:

> I: Out of these three problems, which one was most difficult to answer?
> Earl: … Probably this one [Problem 7]…because it's kind of like, kind of complex, more complex than the other ones, what kinds of adaptations were used by my animal, funnel web spider, to survive in its habitat? It just took longer.

However, a similar improvement pattern was observed in the validity ratio between the two treatments. A considerable improvement in the validity ratio was shown over the first two phases, which may be an indication that an ability to make scientifically valid warrants could be facilitated when students focus solely on making connections between the given data and their claims. However, these data suggest that it is more difficult to develop an ability to weigh and prioritize multiple factors in complicated real-world problems. For example, in Phase III, the H & M students in the fading support group made fewer valid warrants, $M = 1.4$, than those in the consistent support group, $M = 2.1$, while the validity ratios were similar, $M = .79$ - fading support group; $M = .82$ - consistent support group.

H & H profile. Table 5 (c) shows the difficulty rating of the H & H students. Interestingly, all three students in the fading support group agreed that Problem 8 was most difficult and Problem 2 was easiest. All three students mentioned that their difficulty rating was related to how much support was available on the worksheets:

I: Out of three problems which one was the most difficult one?
Matt: This one…third one…[Problem 8]
I: Why is that?
Matt: Because it didn't have the questions, data or evidence questions…
I: So you have to think on your own…
Matt: Yes…
I: Was it difficult to think about questions for what to write about?
Matt: After a while it wasn't that hard to think of it. But at first it was too hard… or it's sort of hard.
I: How about which one was the easiest one?
Matt: Um…probably that one [Problem 2] because it had most support for me.

Despite the lack of scaffolds, the number of valid warrants and the validity ratio were largely unchanged in the explanations over the three phases. However, in the post-treatment explanation, the H & H students with consistent support outperformed the H & H students with fading support in both the number of valid warrants, $M = 2.9$ vs. $M = 1.6$, and the validity ratio, $M = .93$ vs. $M = .83$.

Unlike the other profiles in the fading support treatment, the H & H students recognized the gradual withdrawal of the scaffolds, which means that they were actively using and appreciating the scaffolds provided on the worksheets. In fact, three H & H students in the fading support treatment developed their own questions (Q) on the worksheets when questions were no longer available. One of them was Helen. After realizing the benefits of using questions to formulate her explanations in the previous phases, Helen developed her own questions (see Figure 6). Helen described her experience as follows:

I: I noticed that you have written your own questions under data or evidence.
Helen: Ya.
I: I notice that your questions are not the same as the questions written for other students. How did you come up with these questions?

Helen: I kind of just thought about like what was like…what it would need and I kind of flip back to what does it need to eat and how big its habitat, stuff like that and just kind of wrote those down so that I wouldn't forget them and flip back to them every time I was writing a sentence.
I: So you used these to guide your explanation.
Helen: Ya.

Insert Figure 6 Here

On the other hand, the difficulty rating of the four H & H interviewees in the consistent support group did not match one another. Under the identical support condition, the rating of the H & H students in the consistent support group depended upon the problem complexity originated from their individual investigations. Interview data indicate that the H & H profile in the consistent support treatment used scaffolds not only for constructing explanations but also as a checklist to assess their initial explanations. It is important to recognize that some H & H students were different from students in the other two profiles with regard to their ability to monitor their own learning, as demonstrated by those who developed their own scaffolds when curricular scaffolds were not provided and those who used the scaffolds for other reasons than what was originally intended.

Concluding Thoughts

Scaffolding theory emphasizes gradual and eventual transfer of responsibility to students in learning tasks that are initially beyond their capability. Generalizing fading mechanisms to a broad range of learning situations may be too ambitious a goal to achieve because scaffolding, by definition, responds to the changing needs of individual students. A central issue in characterizing fading mechanisms is to understand roles for students, teachers, and learning materials in promoting the acquisition of expert-like competency in a given inquiry task (Tabak, 2004).

The learning context defined within a distributed cognition environment expands the concept of scaffolding agents to teachers, peers, curriculum materials, and technological artifacts. As a result of these many potential scaffolding agents, it has become increasingly difficult to apply scaffolding systematically across a range of distributed learning environments. Among the four scaffolding features Stone (1998) identifies (using a difficult learning task, providing diverse support features, scaffolding that is adaptive to the learner's progress, and the fading of scaffolds over time), learning environment designers often focus on the first two features: using a difficult task and providing diverse support features (Quintana et al., 2004; Reiser, 2004). Though researchers recognize the importance of adaptive support and fading, these two aspects are often neither part of the design process (Quintana et al., 2004) nor the analysis of scaffolding effects (Sherin, Reiser, & Edelson, 2004). In order to advance the understanding of scaffolding theory, Pea (2004) urges researchers to collect empirical evidence that can justify the differentiation between "scaffolds-with-fading" and "fundamental aides to the doing of science whose fading is unnecessary and unproductive" (p.442).

This study revealed that both consistent- and fading-support treatments helped students gain content knowledge in biodiversity and explanation ability to match given evidence to a claim. The performance difference favoring the consistent support group began to emerge when content-oriented explanation scaffolds started to fade. The comparable gains in content knowledge and explanation ability between the two groups eliminated the possibility that the fading support group somehow learned less content knowledge or lacked the ability to make scientific connections between given evidence and their claim. Rather, the difference in the quality of explanations generated during the intervention period might indicate that the explanation tasks with and without full support did not provide the same kind of opportunity for students to learn how to explain in real-world inquiry contexts. Due to the apparent inconsistency in content knowledge necessary to solve of the range of authentic problems presented here and the added complexity of real-world investigations, fading most likely occurred too soon for most students in this study. However, as these results suggest, the timing of fading can be differentially beneficial for different learners. Early fading might reduce the participation of low ability students while it beneficially allows the most advanced students to create their own scaffolds similar to those that had earlier been modeled for them.

These results do not agree exactly with the study conducted by McNeil et al. (2004) which showed that fading explanation support was actually beneficial for students' scientific reasoning in formulating explanations. The difference between these two studies can be attributed to the different type of explanation scaffolds that were designed to fade in each case. This study used content-rich scaffolds that pinpointed directly what content knowledge to consider in formulating explanations, while the study by McNeil et al. (2004) used content-lean, rhetorical explanation scaffolds. These two studies highlight the point that understanding fading mechanisms is not straightforward and requires a careful examination of the relationships among academic tasks, participants, and scaffolding materials in distributed learning environments.

Results of this study also suggest that scaffolding success depends, perhaps greatly, on adaptive support (Palincsar, 1998). To be effective, scaffolds need to challenge students' current capabilities (Blumenfeld et al., 1987). In Vygotsky's terms (1978), scaffolds should be designed to work within students' Zone of Proximal Development (ZPD). If not, scaffolds can be either too difficult for students to follow or too easy to motivate them. Student interview data indicate that those who actively used and appreciated the explanation scaffolds were students with high explanation abilities. A lack of difference between the two treatment groups in the explanation trajectories suggested that the low explanation ability students did not make the most of the content scaffolds.

This study used achievement data (explanations) as a main indicator to determine the effectiveness of the two scaffolding treatments. However, this study suggests that students' increased metacognitive awareness towards the cognitive process involved in the intended learning be another meaningful indicator (Pea, 2004). For example, three out of seven high explanation ability students in the fading support treatment wrote their own content prompts similar to the questions provided on the worksheets to guide their own explanation building process. Some high ability students in the consistent support treatment used content scaffolds not only for building explanations but for reflecting on their initial explanations. These findings support the importance of recognizing students' metacognition as part of the scaffolding process.

In order to examine the interaction between individual students and explanation scaffolds, this study minimized the impact of teachers and peers. In the distributed learning paradigm, coming to know or learn is not a solitary process. It is important to include all participating human and non-human agents in creating, investigating, and analyzing scaffolding mechanisms. Further research is necessary to illuminate (1) how human agents interact with each other and with physical environments to gain expert-like task competency, (2) how students improve metacognitive awareness of their own learning through scaffolding, (3) how different learning tasks influence design and implementation of scaffolding interventions, (4) how scaffolding interventions can address the multitude of ZPD's in a single classroom, and (5) how one can make a prudent distinction between scaffolds that need to fade and those that need to stay in order to maximize learning outcomes.

References

Anazi, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 64-92). New York: Cambridge University Press.

Anderson, J. R., Boyle, C. F., Corbett, A., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence, 42*, 7-49.

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.

Blumenfeld, P. C., Mergendoller, J. R., & Swarthout, D. W. (1987). Task as a heuristic for understanding student learning and motivation. *Journal of Curriculum Studies, 19*(2), 135-148.

Blumenfeld, P. C., Marx, R. W., Soloway, E., & Krajcik, J. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher, 25*(8), 37-40.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Bruckman, A. (2000). Situated support for learning: Storm's weekend with Rachel. *The Journal of the Learning Sciences, 9*(3), 329-372.

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.

Bybee, R. W. (2000). Teaching science as inquiry. In J. Minstrell & E. H. v. Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20-46). Washington, DC: American Association for the Advancement of Science.

Cavallo, A. M. L., McNeely, J. C., & Morek, E. A. (2003). Eliciting students' understandings of chemical reactions using two forms of essay questions during a learning cycle. *International Journal of Science Education, 25*(5), 583-603.

Cazden, C. B. (2001). *Classroom discourse* (2nd ed.). Portsmouth, NH: Heinemann.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

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

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

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

Chinn, C. A., & Hmelo-Silver, C. E. (2002). Authentic inquiry: Introduction to the special section. *Science Education, 86*, 171-174.

Clement, J. (1991). Nonformal reasoning in experts and in science students: The use of analogies, extreme cases, and physical intuition. In J. Voss & D. Perkins & J. Segan (Eds.), *Informal reasoning and education* (pp. 345-362). Hillsdale, NJ: Erlbaum.

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

Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist, 28*(1), 25-42.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: prompts for reflection in KIE. *International Journal of Science Education, 22*(8), 819-837.

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *The Journal of the Learning Sciences, 12*(1), 91-142.

Doyle, W. (1983). Academic Work. *Review of Educational Research, 53*(2), 159-199.

Edelson, D. C. (1998). Realising authentic science learning through the adaptation of scientific practice. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education*. The Netherlands: Kluwer.

Edelson, D. C. (2001). Learning-for-use: A framework for the technology-supported inquiry activities. *Journal of Research in Science Teaching, 38*(3), 355-385.

Espinosa, R., Dewey, T., Hammond, G., Parr, C. S., Jones, T., Pappas, J., & Myers, P. (2002). *BioKIDS Critter Catalog* (Available at http://www.biokids.umich.edu).

Fleer, M. (1992). Identifying teacher-child interaction which scaffolds scientific thinking in young children. *Science Education, 76*(4), 373-397.

Flick, L. B. (2000). Cognitive scaffolding that fosters scientific inquiry in middle level science. *Journal of Science Teacher Education, 11*(2), 109-219.

Guzdial, M. (1994). Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments, 4*(1), 1-44.

Hempel, C. (1966). *Philosophy of natural science*. Upper Saddle River, NJ: Prentice Hall.

Hodson, D. (1988). Toward a philosophically more valid science curriculum. *Science Education, 72*(1), 19-40.

Hogan, K., Nastasi, B. K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*(4), 379-432.

Keys, C. W. (1999). Language as an indicator of meaning generation: An analysis of middle school students' written discourse about scientific investigations. *Journal of Research in Science Teaching, 36*(9), 1044-1061.

Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.

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.

Klein, P. D. (2000). Elementary students' strategies for writing-to-learn in science. *Cognition and Instruction, 18*(3), 317-348.

Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., & Fredricks, J. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *The Journal of the Learning Sciences, 7*(3&4), 313-350.

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

Kuhn, D. (1992). Piaget's child as scientist. In H. Beilin & P. B. Punfall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 185-208). Hillsdale, NJ: Erlbaum.

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

Kuhn, D. (1997). Constraints or guideposts?: Developmental psychology and science education. *Review of Educational Research, 67*(1), 141-150.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*(4), 495-523.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (Vol. 2). Chicago: The University of Chicago.

Lee, H.-S., & Songer, N. B. (2003). Making authentic science accessible to students. *International Journal of Science Education, 25*(8), 923-948.

Lewis, E. L., & Linn, M. C. (1994). Heat energy and temperature concepts of adolescents, adults, and experts: Implications for curricular improvements. *Journal of Research in Science Teaching, 31*(6), 657-677.

Linn, M. C., & Songer, N. B. (1991). Teaching thermodynamics to middle school students: What are appropriate cognitive demands? *Journal of Research in Science Teaching, 28*, 885-918.

Linn, M. C., & Muilenburg, L. (1996). Creating lifelong science learners: What models from a firm foundation? *Educational Researcher, 25*(5), 18-24.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Lawrence Earbaum Associates.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education, 87,* 517-538.

Lowe, R. K. (1993). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction, 3*, 157-179.

McDonald, S. P. (2004). *Research towards varied models of classroom-based inquiry science.* Paper presented at the Annual Meetings of American Educational Research Association, San Diego, CA.

McNamara, D. S., Kintsch, E., & Songer, N. B. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.

Metz, K. E. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research, 65*(2), 93-127.

Metz, K. E. (1997). On the complex relation between cognitive developmental research and children's science curricula. *Review of Educational Research, 67*(1), 151-163.

Metz, K. E. (2000). Young children's inquiry in biology: Building the knowledge bases to empower independent inquiry. In J. Minstrell & E. H. v. Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 371-404). Washington, DC: American Association for the Advancement of Science.

Millar, R., & Driver, R. (1987). Beyond processes. *Studies in Science Education, 14,* 33-62.

Minstrell, J., & van Zee, E. H. (2000). *Inquiring into inquiry learning and teaching in science.* Washington, DC: American Association for the Advancement of Science.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the National Science Education Standards.* Washington, DC: National Academy Press.

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

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. Annual *Review of Psychology, 49,* 345-375.

Parr, C. S., Jones, T., & Songer, N. B. (2002). *CyberTracker in BioKIDS: Customization of a PDA-based scientific data collection application for inquiry learning.* Paper presented at the Proceedings of the International Conference of Learning Sciences (ICLS), Seattle, WA.

Patterson, E. W. (2001). Structuring the composition process in scientific writing. International *Journal of Science Education, 23*(1), 1-16.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences, 13*(3), 423-451.

Quitana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D. C., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337-386.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences, 13*(3), 273-304.

Resnick, L. B., & Klopfer, L. E. (1989). Toward the thinking curriculum: An overview. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research*. Alexandria, VA: Association for Supervision and Curriculum Development.

Rivard, L. P., & Straw, S. B. (2000). The effect of talk and writing on learning science: An exploratory study. *Science Education, 84*, 566-593.

Sandoval, W. A., & Reiser, B. J. (1997, March 24-28). *Evolving Explanations in High School Biology.* Paper presented at the Annual Meetings of American Educational Research Association, Chicago, IL.

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

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*, 345-372.

Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge-based questioning by children. *Cognition and Instruction, 9*(3), 177-199.

Schunn, C. D., & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science, 23*(3), 337-370.

Schwab, J. J. (1962). The teaching of science as enquiry. In J. J. Schwab & P. F. Brandwein (Eds.), *The teaching of science* (pp. 3-103). Cambridge: Harvard University Press.

Sherin, B., Reiser, B. J., & Edelson, D. C. (2004). Scaffolding analysis: Extending the scaffolding metaphor to learning artifacts. *The Journal of the Learning Sciences, 13*(3), 387-421.

Songer, N. B., Huber, A. E., Adams, K., Chang, H. Y., Lee, H. S., & Jones, T. (2002). *BioKIDS: Kid's Inquiry of Diverse Species, An Eight-Week Inquiry Curriculum using Simple, Powerful Technologies.* The University of Michigan.

Spitulnik, M. W., Stratford, S., Krajcik, J., & Soloway, E. (1998). Using technology to support students' artefact construction in science. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education.* The Netherlands: Kluwer.

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. (2004). Synergy: A complement to emerging patterns of distributed scaffolding. *The Journal of the Learning Sciences, 13*(3), 305-335.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wertsch, J. V., & Stone, C. A. (1985). The concept of internalization in Vygotsky's account of the genesis of higher mental functions. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 162-179). Cambridge, England: Cambridge University Press.

White, B. (1989). *The role of intermediate abstractions in understanding science and mathematics.* Paper presented at the Proceedings of the eleventh annual meeting of the cognitive science society, Hillsdale, NJ.

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

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

Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *The Journal of the Learning Sciences, 12*(2), 145-181.

Table 1

Questions (Q)

	Problem	Questions
1	Among animals I saw in the schoolyard which animals can be grouped together?	• Do animals have internal or external skeletons? • If the animals have external skeletons, how many legs do they have? • If the animals have internal skeletons, do they have feathers or do they hatch from eggs?
2	Are ________ and ______ the same species?	• Do they look alike? • Can they breed with each other? • Do they live in similar places? • Do they behave similarly?
3	How does my invertebrate adapt to the environment where it lives?	• How does the mouth shape help the invertebrate catch food? • How does the shape of legs or wings help the invertebrate move around? • How do antennae help the invertebrate live in where it lives? • How does the color of the body help the invertebrate live in where it lives?
4	Which zone in my schoolyard has the highest animal abundance?	• How many animals were found in this zone compared to other zones? • Does this zone have microhabitats that can afford a large number of animals? • How do microhabitats in this zone afford so many animals? • Were large numbers of the same animals found in this zone?
5	Which zone in my schoolyard has the highest animal richness?	• How many kinds of animals were found in this zone compared to other zones? • Does this zone have microhabitats that can afford different kinds of animals? • How do microhabitats in this zone afford so many different kinds of animals?
6	Which zone in my schoolyard has the highest biodiversity?	• What is the rank of this zone in animal abundance? • What is the rank of this zone in animal richness? • How many animals were found in this zone? • How many kinds of animals were found in this zone? • What kinds of microhabitats exist in this zone to afford animals? • What do these microhabitats offer to animals?
7	What kinds of adaptations are used by my animal to survive in its habitat?	• Are there any microhabitats where your animal is camouflaged? • Does your animal do anything different in very cold or hot weather? • Does your animal have any special body features or behaviors that would help it live there or capture food?
8	Can my animal survive in my schoolyard?	• Does my schoolyard provide food for my animal? • Does my schoolyard provide water for my animal? • Does my schoolyard provide hiding places for my animal? • Does my schoolyard provide spaces to raise young for my animal?
9	What would happen if my animal and my partner's animal met each other?	• Do your animal and your partner's animal compete for food? • Do your animal and your partner's animal compete for shelter? • Can your animal eat or be eaten by your partner's animal?
10	What changes would occur in my food web as a result of a volcanic eruption?	• How does the lack of sunlight affect the habitat environment? • How does the lack of sunlight affect producers in your food web? • How does the lack of sunlight affect consumers in your food web?

Table 2

Exemplars (E)

	Problem	Exemplars
1	Among animals I saw in the schoolyard which animals can be grouped together?	I think a beetle and an ant can be grouped together because they have external skeletons and six legs. These data show that both of them are insects
2	Are ________ and ______ the same species?	I think house mice (*Mus musculus*) and white-footed mice (*Peromyscus leucopus*) are different species because, even though they eat similar kinds of foods and are similar in size, (1) they do not breed with each other; (2) house mice live all over the world while white-footed mice do not occur in some places of US including Florida and west of the Rocky Mountains; (3) house mice breed all year long but white-footed mice breed from March to October if they live in the cold climate.
3	How does my invertebrate adapt to the environment where it lives?	I think slugs live in moist places in fields or woods because their long and skinny body shape plus the absence of legs allow them to move very well around damp soil under wood and rocks. The color of their body is brown, which makes it harder for other predators to find them.
4	Which zone in my schoolyard has the highest animal abundance?	I think Zone B (playground) has the highest animal abundance because 70 animals (playground) were found in Zone B compared to 60 in Zone A and 20 in Zone C. Most of the animals found in this zone were invertebrates that live on trees or under the ground. Trees and soil provide food and hiding place for many invertebrates. In particular, ants (30 ants observed) increased animal abundance greatly in Zone B.
5	Which zone in my schoolyard has the highest animal richness?	I think Zone A (front schoolyard) has the highest animal richness because there were 20 different kinds of animals in Zone A compared to 10 in Zone B and 5 in Zone C. Under the log sitting in the front schoolyard there were at least ten different kinds of invertebrates. The log provides moisture and hiding place for the invertebrates. These different kinds of small invertebrates attracted birds and larger invertebrates.
6	Which zone in my schoolyard has the highest biodiversity?	I think Zone A (front schoolyard) has the highest biodiversity because Zone A was ranked No. 1 in animal richness and No. 2 in animal abundance among all four zones in our schoolyard. In Zone A there were 60 animals in total and 20 different kinds of animals. Zone A has four microhabitats that can be used for animals such as under the log, on plants, in dirt, and in the sky. These microhabitats provide animals with various foods and places for living.
7	What kinds of adaptations are used by my animal to survive in its habitat?	I think my animal (parrot) lives on trees because it has specialized feet with two curling front toes and two curling back toes to help them hang on to branches.
8	Can my animal survive in my schoolyard?	I think my animal (Gray Wolf) cannot live in my schoolyard because Gray Wolves need to have a much larger space than my schoolyard for living and my schoolyard cannot provide enough food, mostly meat, for them.
9	What would happen if my animal and my partner's animal met each other?	I think my animal (mourning dove) and my partner's animal (mosquito) would live together because they do not compete for food and shelter. Mourning doves mostly like to eat seeds on ground, and adult female mosquitoes can feed on mourning doves. Larvae and pupae of mosquitoes live in water and mourning doves live in woodland and forest edges.
10	What changes would occur in my food web as a result of a volcanic eruption?	I think the volcanic eruption would decrease the number of producers such as trees and bushes in our food web because producers cannot survive without sunlight.

Table 3

Sentence Starters (S)

	Problem	Sentence starters
1	Among animals I saw in the schoolyard which animals can be grouped together?	I think ____________ and _____________ can be grouped together because…*[list relevant data or information]*
2	Are ________ and ______ the same species?	I think ______________ and ____________ are the same or different species because…*[list relevant data or information]*
3	How does my invertebrate adapt to the environment where it lives?	I think my invertebrate (name: ______________) lives in __________________ because…*[list relevant data or information]*
4	Which zone in my schoolyard has the highest animal abundance?	I think Zone ______ (zone name:___________) has the highest animal abundance because…*[list relevant data or information]*
5	Which zone in my schoolyard has the highest animal richness?	I think Zone ______ (zone name: ____________) has the highest animal richness because…*[list relevant data or information]*
6	Which zone in my schoolyard has the highest biodiversity?	I think Zone ______(zone name:______________) has the highest biodiversity because…*[list relevant data or information]*
7	What kinds of adaptations are used by my animal to survive in its habitat?	I think my animal lives __________________ because…*[list relevant data or information]*
8	Can my animal survive in my schoolyard?	I think my animal can/cannot (circle one) live in my schoolyard because…*[list relevant data or information]*
9	What would happen if my animal and my partner's animal met each other?	I think my animal and my partner's animal would … because…*[list relevant data or information]*
10	What changes would occur in my food web as a result of a volcanic eruption?	I think the volcanic eruption would … because…*[list relevant data or information]*

Table 4

Two Treatments

Treatment Phase	Problem No.	Two Treatments		Inquiry Problems
		Fading Support	Consistent Support	
Pre-treatment	Pre	none	none	Among animals I saw in the schoolyard which animals can be grouped together?
Phase I	1	Q+E+S	Q+E+S	Among animals I saw in the schoolyard which animals can be grouped together?
	2	Q+E+S	Q+E+S	Are ________ and ______ the same species?
	3	Q+E+S	Q+E+S	How does my invertebrate adapt to the environment where it lives?
Phase II	4	Q+S	Q+E+S	Which zone in my schoolyard has the highest animal abundance?
	5	Q+S	Q+E+S	Which zone in my schoolyard has the highest animal richness?
	6	Q+S	Q+E+S	Which zone in my schoolyard has the highest biodiversity?
	7	Q+S	Q+E+S	What kinds of adaptations are used by my animal to survive in its habitat?
Phase III	8	S	Q+E+S	Can my animal survive in my schoolyard?
	9	S	Q+E+S	What would happen if my animal and my partner's animal met each other?
	10	S	Q+E+S	What changes would occur in my food web as a result of a volcanic eruption?
Post-treatment	Post	none	none	As a zoologist, considering the changes that might occur in your class food web, would you recommend reintroducing Gray Wolves to Michigan? Explain why you think so.

Note. Q = Questions; E = Exemplars; S = Sentence Starters.

Table 5.

Explanation Difficulty Rating of Interviewed Students

(Note. Problem 2 in Phase I; Problem 7 in Phase II; Problem 8 in Phase III)

(a) Low Knowledge and Low Explanation Ability

Intervention	Name	Most difficult problem	Easiest problem
Fading support	Derek	2	8
	Kate (LD)[1]	2	8
Consistent support	Ted	2	8
	Ace (LD)[1]	2	8

Note. [1] LD indicates students with learning disabilities.

(b) High Knowledge and Medium Explanation Ability

Intervention	Name	Most difficult problem	Easiest problem
Fading support	Holly	2	8
	Neo	2	8
Consistent support	Elvis	2	8
	John	8	7
	Terry	2	8
	Earl	7	2

(c) High Knowledge and High Explanation Ability

Intervention	Name	Most difficult problem	Easiest problem
Fading support	Matt	8	2
	Helen	8	2
	Howie	8	2
Consistent support	Lisa	7	2
	Mary	8	2
	Elise	2	8
	Calla	7	8

Figure Caption

Figure 1. A question in claim-evidence test.

Question 3

Some animals are brightly colored and they use their color as a warning to predators. For example, the Monarch butterfly is brightly colored and poisonous if eaten. Predators have learned to leave the Monarch butterfly alone. The Viceroy butterfly is not poisonous, but its color and design are similar to the monarch butterfly as shown below.

A Monarch Butterfly (Poisonous)

A Viceroy Butterfly (Not Poisonous)

What is going to happen to the Viceroy butterfly?

I think a lot of Viceroy butterflies would be eaten / would not be eaten by predators.
Circle one

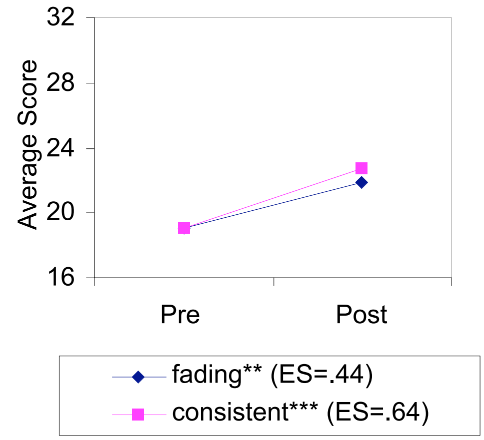
Give reasons that explain your choice.

____

____

Figure Caption

Figure 2. Comparison of content and clam-evidence test results between two treatments.

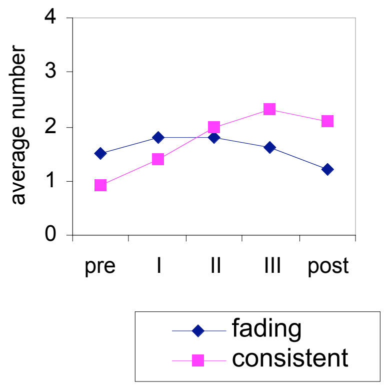


(a) Content Test

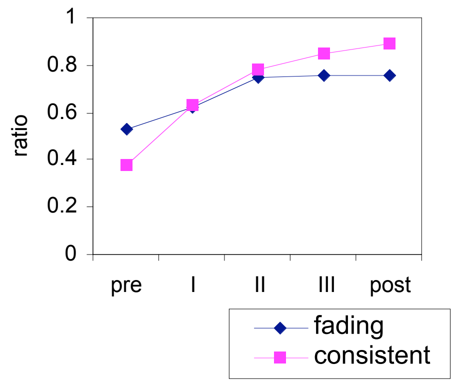
(b) Clam-Evidence Test

Figure Caption

Figure 3. Trajectories of two treatment groups.



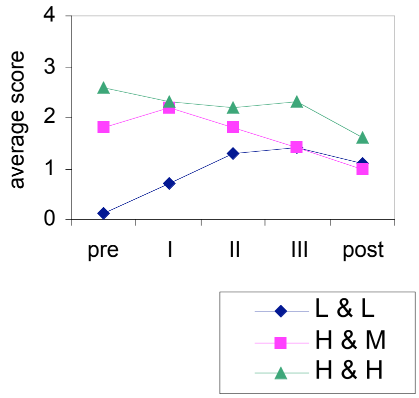
(a) The Number of Valid Warrants



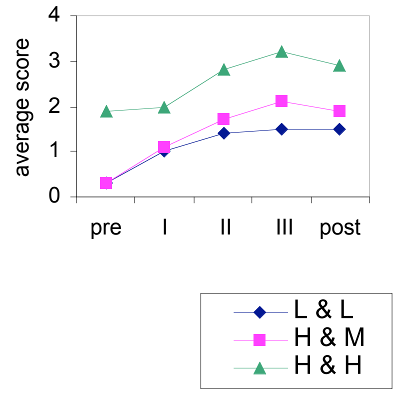
(b) The Validity Ratio

Figure Caption

Figure 4. Trajectories of three learner profiles on number of valid warrants.



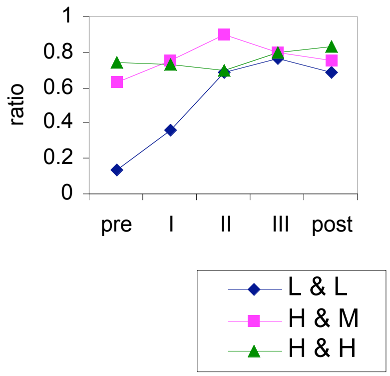
(a) Fading Support



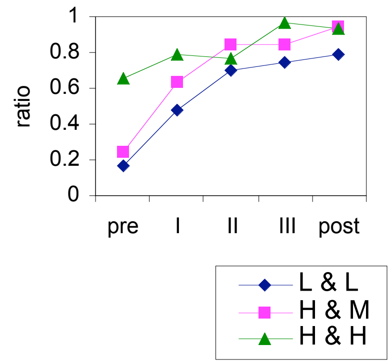
(b) Consistent Support

Figure Caption

Figure 5. Trajectories of three learner profiles on validity ratio.



(a) Fading Support



(b) Consistent Support

Figure Caption

Figure 6. Helen's explanation about Problem 8 (Phase III).

Question: Can my animal (name: Coyote) live in my schoolyard?

	Your Explanation
Claim	I think my animal can/cannot (circle one) live in my schoolyard
Data or Evidence • Is their food soarce there? • Could they raise young there? • Is it big enough? • Are their predators there?	because…*[list relevant data or information]* Except for garbage and some insects, none of their prey lives there. The woods are too small to give prey or raise young and the whole area is too small. Also, humans, o of their predatori, Lives the It would be a dangerous live for a coyote Pack in the school yard.