

Explanation-Driven Inquiry: Integrating Conceptual and Epistemic Scaffolds for Scientific Inquiry

WILLIAM A. SANDOVAL

*Graduate School of Education and Information Studies, University of California,
Los Angeles, CA 90095-1521, USA*

BRIAN J. REISER

School of Education and Social Policy, Northwestern University, Evanston, IL 60208, USA

Received 23 July 2002; revised 19 May 2003; accepted 22 May 2003

ABSTRACT: Science education reforms consistently maintain the goal that students develop an understanding of the nature of science, including both the nature of scientific knowledge and methods for making it. This paper articulates a framework for scaffolding epistemic aspects of inquiry that can help students understand inquiry processes in relation to the kinds of knowledge such processes can produce. This framework underlies the design of a technology-supported inquiry curriculum for evolution and natural selection that focuses students on constructing and evaluating scientific explanations for natural phenomena. The design has been refined through cycles of implementation, analysis, and revision that have documented the epistemic practices students engage in during inquiry, indicate ways in which designed tools support students' work, and suggest necessary additional social scaffolds. These findings suggest that epistemic tools can play a unique role in supporting students' inquiry, and a fruitful means for studying students' scientific epistemologies.

© 2004 Wiley Periodicals, Inc. *Sci Ed* **88**:345–372, 2004; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/sce.10130

INTRODUCTION

A primary goal for current inquiry-based reforms in science education is that students develop an understanding of the nature of science by doing science (NRC, 1996; Rutherford & Ahlgren, 1990). Simply engaging in inquiry, however, is not enough to develop students' ideas about the nature of science (Lederman, Wade, & Bell, 1998). Our view is that inquiry-based reform efforts must emphasize that the processes scientists value for generating and validating knowledge emerge from epistemological commitments to what counts as scientific knowledge. Scientific inquiry is generally defined as a process of asking questions, generating data through systematic observation or experimentation, interpreting data, and drawing conclusions (e.g., NRC, 1996; White & Frederiksen, 1998). The epistemic aspects

Correspondence to: William A. Sandoval; e-mail: sandoval@gseis.ucla.edu
Contract grant sponsor: James S. McDonnell Foundation.

of inquiry include knowledge of the kinds of questions that can be answered through inquiry, the kinds of methods that are accepted within disciplines for generating data, and standards for what count as legitimate interpretations of data, including explanations, models, and theories. Placing these epistemic aspects of scientific practice in the foreground of inquiry may help students to understand and better conduct inquiry, as well as provide a context to overtly examine the epistemological commitments underlying it. Our research and design efforts have explored how we can design learning environments that make epistemological goals for scientific work visible in ways that support students' inquiry. Rather than focus simply on the processes of inquiry that students might engage in, we have focused students' inquiry on the kinds of products those processes are intended to construct and evaluate. Our goal through this work is to both design and enact effective learning environments in science classrooms while using these classrooms as sites for research into students' understanding of scientific practice.

This paper describes an approach to providing epistemic support for inquiry learning. First, we present a conceptual framework that describes how epistemological commitments to kinds of scientific knowledge influence processes of scientific inquiry. Second, we describe how this framework guided the design of a technology-supported learning environment for evolution and natural selection that integrates epistemic and conceptual scaffolds for inquiry. We then present evidence from studies of the learning environment in initial classroom use to explore the interplay of epistemic scaffolds and students' decision making during inquiry. Our designs have been refined through multiple cycles of development and classroom use. We describe how the data have informed our redesign efforts and consider what these classroom trials suggest about the viability of the approach, following a design-based research methodology (Brown, 1992; Collins, 1992; Design-Based Research Collective, 2003).

FRAMEWORK: SCIENCE AS EPISTEMIC PRACTICE

What kind of science should children learn in school? Typical science instruction teaches what Duschl (1990) has called "final form science," where theoretical ideas are presented as incontrovertible facts, stripped of the history of their development. Such an approach effectively removes students as producers of scientific knowledge, locating authority for the thematic content of science and appropriate argumentation strategies with teachers (Lemke, 1990) and textbooks. Typical science instruction leads students to develop, not surprisingly, an undesirably naïve view that science is an unproblematic accumulation of facts that describe the world (Carey et al., 1989; Driver et al., 1996; Lederman, 1992). Students see that the best way to learn such facts are to pursue shallow learning strategies, such as memorization and rote application of procedures and formulas, that do not promote deep learning (Hammer, 1994; Linn & Songer, 1993). Very few students develop epistemological views of science as a process of building and revising models and theories about the world, rather than the discovery of facts in the world (Driver et al., 1996).

This is clearly a challenge for inquiry-oriented approaches to science instruction, since these involve students explicitly in theory/model building and revision. Students' difficulties in conducting inquiry are well-documented. Students have difficulty formulating appropriate research questions and plans to investigate them (Krajcik et al., 1998). Students have considerable difficulty in designing experiments and interpreting results (for a review, see Zimmerman, 2000). Such difficulties stem in part from students' lack of knowledge about particular domains. When students have more sophisticated conceptual knowledge in a domain, they investigate it more systematically (Schauble et al., 1991). Students also seem to have oversimplified models of the purpose of experiments. Schauble and colleagues,

for example, found that sixth grade students typically did not see experiments as efforts to isolate causal variables. After instruction on the purpose of experimentation, however, these students were able to design better experiments (Schauble et al., 1995). Similarly, Dunbar (1993) found that undergraduate subjects instructed to explain data, rather than verify a given hypothesis, were more systematic and effective in their exploration of a computer microworld to discover genetic function. These efforts to focus students on causal explanation suggest the power in making the epistemic goals of scientific inquiry explicit for students.

Studies of students' strategic reasoning during science investigations and of their professed beliefs about the nature of science have occurred largely independently from one another. A few studies suggest that students with more constructivist epistemological frameworks, who see scientific knowledge as both constructed and changeable over time, perform better in inquiry contexts (Linn & Songer, 1993; Tobin, Tippins, & Hook, 1995; Windschitl & Andre, 1998). Kitchener (1983) argued that what she called epistemic cognition acts as a control over metacognition during ill-structured problem solving, such as open-ended inquiry. That is, knowledge about what needs checking, what counts as "being done," and so forth, makes meaningful metacognitive monitoring possible. Although the above work shows associations between students' epistemological perspectives and their inquiry strategies, there seems to be little evidence that conducting inquiry, in itself, changes students' conceptions of the nature of science (Lederman, Wade, & Bell, 1998). Calls for making argumentation a central practice of science instruction presume that making the epistemic aspects of inquiry more explicitly central may improve students' abilities to conduct inquiry and support their epistemological development (Driver, Newton, & Osborne, 2000; Kuhn, 1993).

We approach inquiry instruction as a cognitive apprenticeship (Collins, Brown, & Newman, 1989) into scientific practice. This approach has two important consequences. First, when we consider what it might mean to engage students in "authentic" scientific inquiry, we focus on engaging students in the reasoning and discursive practices of scientists, and not necessarily the exact activities of professional scientists (Reiser et al., 2001). Second, viewing such reasoning practices as a form of apprenticeship emphasizes the epistemic aspects of scientific practice. For science, important practices include being able to formulate researchable questions, design and conduct informative investigations, and formulate persuasive arguments. Judgments of "researchable," "informative," and "persuasive," however, are grounded in epistemological terms, they have to do with disciplinary values about what counts as scientific knowledge and what methods can satisfactorily generate such knowledge. Such values change across disciplines, and within disciplines over time, as the questions of interest and the methods used to answer them change (e.g., Kuhn, 1970; Mayr, 1988). A deep understanding of science, and an ability to really conduct inquiry, thus demands an understanding of the epistemic aspects of the practices of specific scientific disciplines. Our view has been strongly influenced by the cognitive apprenticeship model and ideas of situated learning (Lave & Wenger, 1991), and is consistent with recent educational perspectives on scientific practice drawn from science and technology studies (Kelly, Chen, & Crawford, 1998a; Roth & McGinn, 1998).

We have developed a conceptual framework to articulate the relations between general scientific epistemological commitments and disciplinary paradigms. This framework emphasizes the influence of epistemological commitments on strategies for pursuing inquiry, and the dialectic relation between disciplinary knowledge and epistemological commitments. 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. In fact, controlling variables across experiments is just one form of a broader strategy of systematic comparison to isolate causal relations. Such systematicity is pursued because it leads to valued forms of knowledge.

Knowing that controlling variables is important does not help without some idea of what the relevant variables might be and how to consider their possible relations. Discipline-specific knowledge is needed to guide successful inquiry (Schauble et al., 1991; Tabak et al., 1996). Specific theories suggest investigative strategies for applying theory to specific cases. For example, the theory of natural selection argues that environmental factors select for specific traits in organisms. Explaining a case of natural selection therefore entails documenting systematic differences in traits between individuals in a population, and relating them to factors in the environment. In this way, central theories in a discipline provide explanatory frameworks that guide the implementation of more general epistemic goals. Multiple sources of disciplinary knowledge might be needed to explain any particular phenomenon. Knowledge of specific ecosystems is needed to judge what environmental factors might matter for a particular organism, for example, and how change in that factor can be documented and related to the differential survival of individuals in a population. As Collins et al. (1989) point out, the expert strategic and epistemic knowledge needed to produce and evaluate causal explanations is usually not made explicit in instruction. Our question, then, is how to make such knowledge explicit in ways that students can use during their own inquiry.

DESIGN PRINCIPLES FOR SUPPORTING EXPLANATION-DRIVEN INQUIRY

The nature of the tools that people use have a strong influence on how they think about accomplishing tasks with those tools (Norman, 1988; Pea, 1993). Cognitive tools can help guide students' thinking in productive ways (Lajoie, 1993; Reiser, 2002). We have used the view of scientific practice just described to generate design principles for the development of learning environments to support high school students' inquiry into complex problems in biology. Our main effort has been to develop software tools that represent key epistemic criteria in discipline-specific forms, and to structure students' activities around these tools in ways that encourage valued scientific practices and scaffold students' investigations. We build from the notion of epistemic forms (Collins & Ferguson, 1993), that forms for representing knowledge constrain the strategies used to make and reason with that knowledge. Our description here focuses on a particular tool, ExplanationConstructor, designed to support students' construction and evaluation of explanations through their inquiry, and the activities that surround students' use of the tool. Explanation-driven inquiry entails a shift both in the nature of students' work in the classroom and their underlying view of that work. Accomplishing this shift requires tools that shape the ways that students construct the products of their work, curricular activities that emphasize the valued criteria of these products, and teaching practices that support students' understanding of these criteria and help to connect their inquiry experiences to core disciplinary theories.

Each of the following sections describes a key principle underlying our design. For us, as designers, each principle represents an aspect of our approach to supporting inquiry and explanation; each is something that we have implemented and is reified within ExplanationConstructor and the activity structures we have designed around it. As

researchers, each principle is better seen as a conjecture about how to support learning. They suggest those aspects of students' inquiry practices we need to examine to assess the possible influence of the design on students' learning. This description exemplifies how our conceptual framework gets reified in design and provides the context for exploring how the design functions in classrooms.

Grounding Process in Products: Explanation-Driven Inquiry

The overarching principle guiding our approach is to frame students' inquiry as an effort to explain phenomena, emphasizing the form that students' explanations should take. This focus on the form of the product is intended to help students attend to epistemic features of explanations, and provide epistemic goals for their inquiry. The design decisions reflected in ExplanationConstructor and the activities during which it is used emphasize two criteria for explanations in particular: (a) the articulation of coherent, causal accounts and (b) the use of data to support causal claims. These criteria are core epistemological goals for scientific explanations, although they are not the only ones of potential interest (cf. Gitomer & Duschl, 1995). We have developed three computer-based investigation environments covering natural selection among finches in the Galápagos islands, the evolution of bacterial resistance to antibiotics, and the ecology of panthers in North America.¹ Each environment poses a question that students work collaboratively to explain. These programs and the broader curricula we have developed around them are described in more detail elsewhere (Reiser et al., 2001).

We developed ExplanationConstructor to organize students' investigations of these problems and structure their efforts to construct explanations. ExplanationConstructor is an electronic journal in which students build an ongoing record of their investigations. The most recent version of ExplanationConstructor is shown in Figure 1, with an example of actual student work. ExplanationConstructor reflects a general strategy to create representational tools that make explicit important conceptions and processes of scientific inquiry within a discipline. Students' investigations are framed around the desired products, with the expectation that this will encourage better investigation and better learning.

Link Explanations to Specific Questions

The kinds of explanations that scientists generate about specific phenomena are constrained by the questions they ask of them. Most formulations of scientific inquiry frame the process as starting with particular questions (e.g., NRC, 1996; Rutherford & Ahlgren, 1990; White & Frederiksen, 1998). Explanations are answers to particular questions, and this connection is important epistemically. The evaluation of the worth of any explanation is in relation to its value as an answer to the original question. Even when new explanations fundamentally change the questions that scientists pursue, they still originate in efforts to answer particular questions.

In our curricula, students' inquiry is posed as an effort to answer particular questions. Students investigate these questions by exploring a particular investigation environment (an example is given below). Questions are framed to emphasize causal explanation, and students are encouraged by their teachers to use evidence to support their explanations. Using ExplanationConstructor, students organize their investigations around the questions they are trying to answer (upper left in Figure 1). The program allows students to organize

¹Available through the Center for Learning Technologies for Urban Schools (www.letus.org/bguile/) or through the BioQuest consortium (www.bioquest.org).

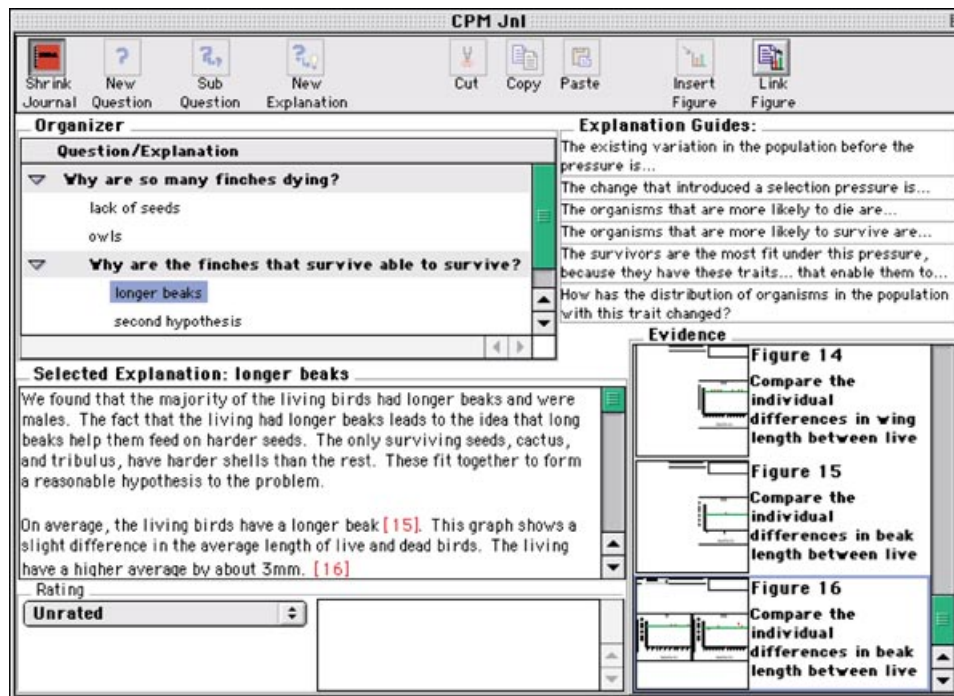


Figure 1. ExplanationConstructor v2.0, showing the final journal of a group of high school students investigating natural selection using the Galápagos Finches program. (See text for descriptions of the features.)

questions hierarchically; a new question can be entered as a sub-question to a larger question. This feature is intended to support students' decomposition of complex questions into more manageable ones. New explanations that students create are always connected to selected questions (Figure 1).

Other tools also structure students' investigations around questions and candidate answers to them, in one form or another (e.g., O'Neill & Gomez, 1994; Scardamalia & Bereiter, 1993–1994). Scardamalia and Bereiter argue that this way of structuring discourse supports students' intentional learning. We build on this notion of intention to encourage students to focus on the epistemic aspects of their inquiry within specific investigative contexts. Framing investigation as the production of explanations for questions is our way of making explicit the link between epistemological commitments to forms of knowledge and investigative strategies. This link should support students' monitoring of their own investigations. Students' recorded questions are persistent reminders of their investigative goals, and a dynamic record of the changing course of their investigation. Because explanations are linked directly to specific questions, this provides a simple way to monitor whether or not questions have been answered.

Represent Theories as Explanatory Frameworks

A key feature of ExplanationConstructor is that it provides discipline-specific scaffolds to guide students' construction and evaluation of their explanations. *Explanation guides* represent explanatory frameworks as a set of connected prompts that highlight their conceptual content and their epistemic structure. Conceptually, explanation guides focus students on the possible content of explanations. In Figure 1, for example, the shown guide visually

prompts students for the important components of a natural selection explanation. This conceptual guidance is motivated directly from our framework, to provide two primary benefits. First, because explanation guides are discipline-specific they have the potential to suggest specific investigative actions that students can take. By providing students the causal components of the target product of their inquiry, we may effectively guide their investigative processes. The second benefit is that explanation guides provide a concrete means for students to monitor their progress. They have finished constructing an explanation only when they have written an explanation that answers each prompt in a guide.

In concert with strategic scaffolds within particular investigation environments, explanation guides can also emphasize important conceptual hurdles in the domain. For example, students typically fail to notice the importance of individual variations to the process of natural selection, and often view evolution as the adaptation of individuals rather than populations (Bishop & Anderson, 1990; Brumby, 1984; Greene, 1990; Settlege, 1994). Consequently, the “selective pressure” guide shown in Figure 1 specifically frames explanations in terms of the effects of environmental pressures on individuals, on the specific traits of individuals, and advantages from those traits. We have created explanation guides for specific theoretical frameworks to focus students on such difficult conceptual ideas as they explain particular problems. We have developed guides in consultation with professional biologists and teachers, with an aim toward faithfully representing the key aspects of scientific theories in a language that high school students can understand and use.

Epistemically, explanation guides encourage students to think about theories as explanatory frameworks, superordinate to explanations for specific events. Each problem that students investigate includes multiple explanation guides to choose from. Because students have to choose guides for each explanation, they are encouraged to map their emerging understanding into domain theory, to place their ideas within a particular explanatory framework. Also, explanation guides make clear that these frameworks include distinct but related components that must cohere to make a good scientific explanation. An environmental pressure, for instance, has to be connected to individual differences in traits and the advantages of a particular trait. This provides a context for students to consider the relationships between theory, specific explanations based on a theory, and the data that may support a particular explanation.

By integrating conceptual and epistemic scaffolds in this way, explanation guides make explicit for students the theoretical frameworks they can use to direct their investigation and analyze their progress toward a satisfactory explanation. Students should come to see that a convincing scientific argument for the specific problem needs to be framed within this organizing structure. Thus, the contents of the explanation guide prompts focus students’ attention on key conceptual ideas by making them explicitly problematic (Reiser, 2002).

Link Evidence to Causal Claims

The major work of scientific explanation is to coordinate patterns of data with causal claims about what the data mean. Data thus have quite a different epistemological status from the causal claims derived from them. This distinction often is not made by students. Instead they seem to view explanations as being embodied in data, not interpretations given to data (Carey & Smith, 1993; Kuhn, 1993; Kuhn, Amsel, & O’Loughlin, 1988). The authoritative discourse of typical science instruction encourages this overobjectification of data (Lemke, 1990), while ignoring the historical development of theoretical ideas (Duschl, 1990). Using ExplanationConstructor, students have to select specific pieces of data as evidence, and then link them to specific causal claims (see Figure 1). Thus, the distinction between claim and

evidence is made both in the representations students use and in their manipulations of those representations. This is a way to emphasize the influence that disciplinary knowledge has on investigation strategies—such strategies aim to generate evidence for or against knowledge claims. We point out that the tool itself does not necessarily encourage students to see data less objectively. Rather, the creation of potentially competing artifacts provides a context for the meaning of the data to be argued about and negotiated by the class.

Structured Opportunities for Epistemic Reflection

Our design for new practices of science in the classroom includes both tools and tasks that exploit features of those tools to promote student talk about target epistemic ideas. Our goal is to help students engage in a set of epistemic practices around the construction and evaluation of scientific explanations. Drawing on the role of reflection in cognitive apprenticeship (Collins, Brown, & Newman, 1989), we have designed activities in which groups critique each others' explanations, and self-assess their own progress. These activities are designed to reinforce the goals of the explanation task; they focus students on the causal coherence of their claims, their fit with available data, and with overarching disciplinary theories. We have defined two opportunities for reflection to give students a chance to revise and improve their explanations and to work against a notion that they construct "final" answers to these questions. Our experiences in the classroom have led us to move toward designing more structure for these discussions. We will explain the reasons why as we summarize our initial classroom studies below.

EXPLANATION-DRIVEN INQUIRY IN THE CLASSROOM

We present case study and other data from two classroom implementations focused specifically on understanding the role that ExplanationConstructor plays in supporting students' efforts to construct and evaluate explanations. Elsewhere we have reported on the explanations that students constructed during these implementations (Sandoval, 2001, 2003). Following from our model of epistemic reasoning outlined above, ExplanationConstructor and the associated material and task scaffolds ought to provide support for students to make strategic decisions during complex investigations. We expect such support to manifest itself in several ways. Because the prompts in ExplanationConstructor are discipline-specific, they ought to suggest to students the kind of data they need to collect to be able to explain particular aspects of a problem (e.g., data that show an environmental pressure, or individual differences in survival). We also expect that the combination of representations ought to help students monitor their progress in epistemically important ways. The explicit links between questions, explanations, and data enable students to assess whether they have answered their current question, completed their current explanation, and provided evidence to support their claims.

In the following sections we describe two initial classroom studies that explore the inquiry practices that students engage in when using our design and how the design is implicated in these practices. We have followed a design-based research approach where studies of the design in use inform changes to the design and revise our understanding of the underlying learning issues (e.g., Design-Based Research Collective, 2003). Our first study sought to understand the practices of students' inquiry and the ways in which ExplanationConstructor did or did not support that inquiry. Our findings from this study motivated several revisions to our design. In our second study, we shifted our focus from students' construction of explanation to explanation evaluation.

Year 1: Connecting Student Practices to Tool Use

Setting. The first version of our evolution curriculum, including ExplanationConstructor, was implemented in the Spring of 1997, in a school in an affluent, predominantly White suburb of Chicago. We chose this school for this study for two reasons. A teacher there was interested in our work and shared our goals for engaging students in inquiry, and was willing to collaborate to integrate our designs into his existing curriculum. Second, this school already had the technological infrastructure to support students' classroom work. Mr. Gray (a pseudonym) had taught high school biology and chemistry for 15 years, and had bachelors and masters degrees in biology. In 1997, we implemented our evolution curriculum in two regular level and one honors level introductory, ninth-grade biology courses. Students earned placements into the honors course through a written exam unrelated to biology; so these students may have been better writers or possibly more motivated, but were not more knowledgeable about biology. In the end, we found no important differences in performance between any of the three classes (Sandoval, 2003). There were a total of 69 students in these classes (42 girls, 27 boys).

The BGulle Evolution Unit. The evolution unit lasted 4 weeks in each class, as outlined in Table 1. Mr. Gray introduced Darwin's theory of evolution with a lecture describing the major ideas. This was followed by lab activities in which students documented, graphed, and discussed patterns of individual variation in several organisms, including themselves (e.g., hand widths). The major purpose of these labs was to help students notice individual variation, to consider possible effects from such variation, and to graph patterns of variation in ways that would help them understand the data in the computer-based investigation environments. During the second week, students worked in small groups to explore the Galápagos Finches (GF) environment. Following the GF investigation, students were involved in a variety of activities, including a natural selection simulation activity in which students worked in pairs to "hunt" colored dots on multicolored backgrounds, watching animated videos of evolution, and examining aspects of human evolution through textbook lab activities (e.g., comparing cranial capacity in humans and chimps). In the third week, students returned to the computer to collaboratively explore TBLab, a program we developed that allows students to explore how the bacteria that cause tuberculosis evolve resistance to antibiotics. Finally, the last few days of the unit were taken up by whole-class discussions in which Mr. Gray elicited students' syntheses of the common aspects of the investigations they had conducted during the unit and mapped them into a common framework of Darwinian natural selection.

Exploring Epistemic Practice. In this first classroom study, our main interest was to document students' practices of investigation and explanation, and explore how explanation guides and the explicit representation of epistemically distinct objects (i.e., questions, explanations, and evidence) may have supported those practices. Both the GF and TB problems posed overarching questions to students, which they investigated in groups of three (rarely four). We wanted to know if students posed their own questions, and when. When did they decide that they needed to or could propose an explanation? How did they decide what data to look at? How did they decide that particular data were relevant evidence for claims? How did they negotiate the content of their questions and explanations? What role did ExplanationConstructor play in students' discussions?

To explore students' epistemic practices and ExplanationConstructor's possible role in them, we recorded groups of students as they collaborated to investigate and explain both the GF and TB problems in the unit. As will become clear below, these investigations were quite

TABLE 1
Outline of Evolution Unit as Implemented in Year 1

Day of Unit	Activity	Purpose	Developed By
1	Darwin lecture	Introduce theory and historical context	Teacher
2–3	Variation labs	Practice observing and graphing individual variation	Teacher and researchers
4–8	Galápagos finches	Investigate and explain natural selection, especially relate individual differences to changes in the population	Researchers
9–10	Colored dots simulation	Reinforce selection as a population effect; show preexisting variation	Teacher and researchers
11–12	Human evolution labs	Demonstrate human genetic lineage; make unit relevant to students	Teacher
13–17	TBLab	Investigate and explain natural selection; reinforce importance of preexisting variation, and the genetic basis of variation	Researchers
18–20	Wrap-up discussions	Develop thematic coherence across experiences	Teacher and researchers

Year 2 implementation dropped colored dots simulation and spent more time on explanation evaluation.

open-ended and students decided for themselves when to record questions and explanations, and when to link evidence. We examined when they made such decisions, the nature of each decision, whether or not they used ExplanationConstructor as they negotiated decisions, and how they did so. In each class, groups mostly self-selected, with Mr. Gray maintaining the power to veto groups he thought would not work well together. We videotaped one group from each class during each investigation, resulting in six video case studies. There were approximately 5 h of videotape for each group. Unfortunately, two cases from the TB investigation could not be analyzed in detail because of poor audio quality. Videotape data were augmented by audiotapes of four other groups and researcher field notes.

Guided by methods of interaction analysis (Erickson, 1992; Jordan & Henderson, 1995), we examined videos of particular groups' investigations across several class periods, logged significant episodes, and then focused our analytic efforts on identifying and analyzing specific events of epistemic discourse. We defined as significant those episodes in students' collaboration where they articulated new questions to investigate, proposed a new or refined an existing explanation, changed their investigative tack, or chose data to use as evidence. Our goal was to figure out where strategic decisions were being made and the role ExplanationConstructor played, if any, in guiding those decisions.

We began our video analyses by arbitrarily selecting one of the cases and exploring that case in depth to generate categories of the kinds of practices students engaged in. We then used these initial categories as lenses to examine three other video cases, looking for confirming or disconfirming evidence for each, as well as for potentially new categories. Our four audio-only cases were of some use in corroborating the presence of epistemic talk, but without a visual reference to what students were looking at as they talked these data could not support claims about tool use. We present four kinds of epistemic practices that were corroborated across multiple video cases and occurred multiple times in each and in which ExplanationConstructor appeared to play an important role. These are (a) epistemically oriented monitoring, (b) planful investigation, (c) negotiating explanations, and (d) evidence evaluation. A fifth important epistemic practice in these investigations was recognizing important data, which we saw occur largely within the investigation environments rather than ExplanationConstructor.

We have chosen to present these kinds of practices through a single case, to provide a sense for the points in an investigation where students focus on particular epistemic aspects of their investigation. The case we have chosen is the first case we analyzed in depth and where our analytic themes were developed. This is a strategic choice to enable readers to develop a sense of the nature of the inquiry students were engaged in, although it comes at the cost of presenting the breadth of practices across cases. We strive to point out below how common each of our identified themes were across cases. Our example case included two girls, Franny and Janie, and a boy, Evan (pseudonyms), working together on the GF problem. The GF problem given to students is this: “In 1977 the population of medium ground finches on the small Galápagos island of Daphne Major declined by more than 60 percent. Why did so many birds die that year? How were the surviving birds able to survive?” Together, these questions frame a problem of natural selection, although the problem is not explicitly framed that way for students. Mr. Gray told the class they must explain *how* and *why* some birds survive while other birds die, emphasizing the demand for causal accounts. He also told them they would be graded on their use of data as evidence, although there was no discussion of what it meant to “use” data.

Epistemically Oriented Monitoring of Progress. Within such an open-ended inquiry task as the GF problem, keeping track of one’s progress is essential to success. Students need to be able to maintain a sense of where they are in their investigation, what particular part of the problem they are exploring, and how close they are to having a satisfactory answer. Students often lose track of where they are in a series of experiments (Kuhn, Schauble, & Garcia-Mila, 1992; Schauble et al., 1991). Evan, Franny, and Janie and other groups in these classes frequently paused to check on their progress, and relied on features of ExplanationConstructor to help them do so in epistemically productive ways.

In the following excerpt, Evan exploits the prompts (Figure 2, showing an earlier version of ExplanationConstructor) to remind the group of where they are at in their current explanation, and what they need to do next.² The group had discovered a graph showing rainfall in the wet and dry seasons for several years. They noticed that in the wet season of 1977 there was very little rain, much lower than usual (the picture to the right and behind the front window in Figure 2). They decided to start a new explanation, that they titled “Rainfall,” and described their findings (the top text box in Figure 2). In this earlier version

²In all excerpts, student speech is presented verbatim with invented spellings to illustrate idiomatic speech. Each turn of student talk starts on a new line. Points at which students interrupt each other and their speech overlaps are indicated by double slashes, //. Short pauses are indicated by dashes. Comments inserted to indicate gestures or to clarify the context of students’ remarks are enclosed in brackets [like this].

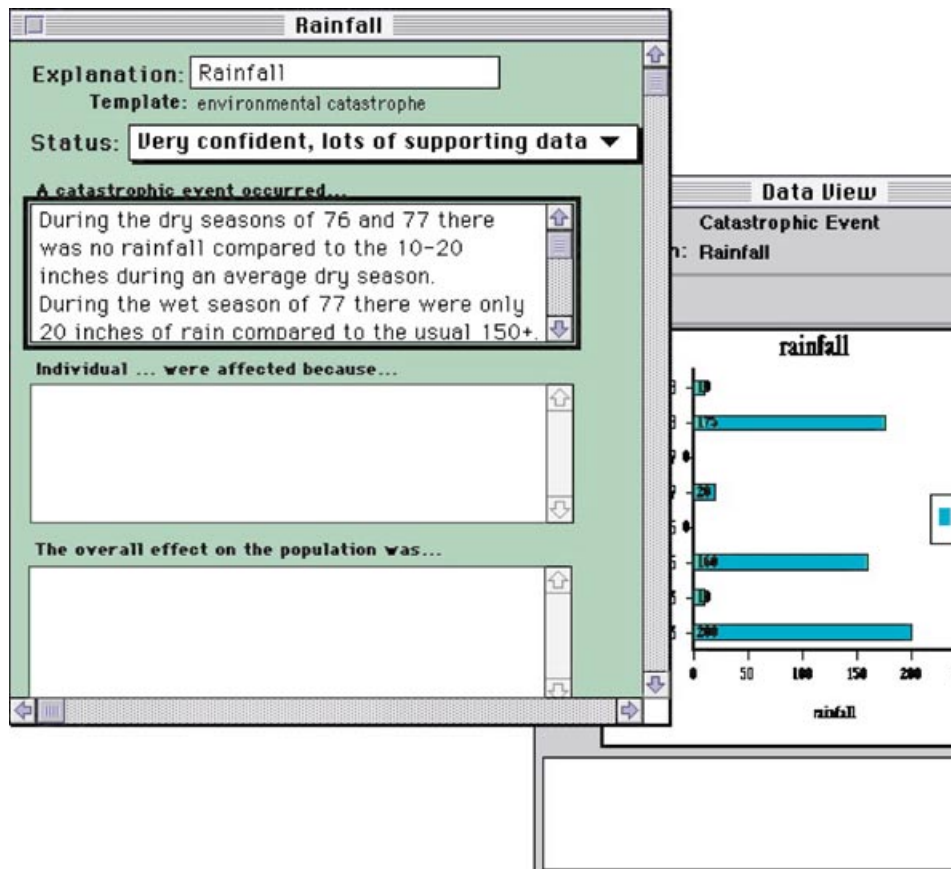


Figure 2. Students use explanation guide prompt to monitor progress.

of ExplanationConstructor, explanation guides were represented as separate “templates,” where the guide prompts were integrated with text boxes for students’ to write in. (We explain later our reasons for changing this representation.)

After typing in the text for the first component, Janie poses a question to her partners:

Janie: So, now where do we want to go?

Franny: You guys, we need sub-questions.

Evan: No, we’re still answering that question [pointing to question in journal].

Franny: We are?

Evan: Yeah.

Janie: We are?

Evan: Yeah, we have to find out which individuals were affected. [pointing to 2nd, still blank component of explanation]

Note that while Janie’s initial question is just a general kind of progress check, Evan’s use of the guide prompt focuses this check in explicitly epistemic terms. He draws the group’s attention to their explanation, pointing out that it is unfinished—they haven’t explained

how the lack of rain has affected individuals. The guide prompts do more than suggest in a general way that their current explanation is unfinished—they directly suggest what the group must do to complete it. This, in turn, suggests the kind of data they need to look for next: data that can explain how individuals were affected by the lack of rain. The need to organize their findings in terms of an explicit framework helps the group attend to unresolved components of their explanations.

From our observations, this group was highly attentive to this sort of monitoring. Other groups that we analyzed varied in the amount of overt monitoring of their own progress, but all of the groups we analyzed referred explicitly to elements in their journals when they did so. Attention to these features was an explicit part of the task as framed by the teacher and was supported by the software. We emphasize that such monitoring focused students' conversations in epistemic terms, but did not necessarily enable all groups to have equal success in solving this problem.

Planful Investigation. Schauble and colleagues (1991) used the term *planful* to refer to investigation strategies that were driven by students' plans or goals for their investigation. We distinguish planful investigation from group monitoring to emphasize the forward-looking nature of planning. Such looking ahead is encouraged by, and may rely upon, monitoring current progress. In fact, as the previous and next example show, Evan, Franny, and Janie often used their reflections on their progress to generate ideas about next steps. It was not that these students or any of the other groups we observed devised plans ahead of time and then carried them out. Instead, we observed a fairly frequent back-and-forth between browsing data in the investigation environments and planning investigative steps based on current progress. That is, plans changed in response to new data and the effort to figure it out.

In the next excerpt the group has completed their explanation for why the finches died (in Figure 3), and are pondering their next move. Their deliberations are centered around the creation of a new question in their ExplanationConstructor journal. The effort to organize and record their progress encourages the group's talk about what to do next.

Janie: Now we have to figure out—

Evan: Now we have to write another question.

Franny: Yeah, we need another question. We have one sub-question.

Evan: Actually, you know—cuz we know that they all died, that this—this—[gesturing to components of "Rainfall" explanation], but we don't know that they're correlated. We know that they all died those years//

Janie: //Does the fact that there are no plants, affect the—? [typing this question]

Evan: Are the deaths of the plants—

Franny: —the reason why. Yeah. But how are we gonna answer that? Where are we gonna go look?

Janie: It's the lack of seed.

We saw a lot of variability between groups in the amount of explicit planning of next steps. In all of the groups we observed, this happened at least some of the time. Again, it is almost a built-in demand of the task that as students construct an explanation they take stock of what they know and what they still need to find out. It is important to note, therefore,

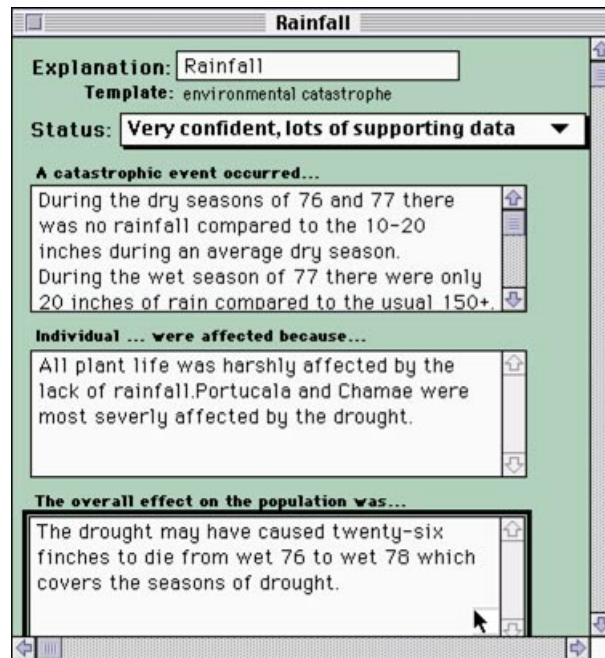


Figure 3. Evan, Franny, and Janie's explanation for the finches death. Upon completing this, the group pauses and decides what to do next.

that the software itself does not provoke such conversations out of thin air, but supports them to the extent that representational features of the tool are consistent with students' understanding of the task. With Janie and her peers, once they have chosen to record a new question, they need to agree on what the question is and this leads to a consideration of how they might answer it.

Negotiating the Terms of an Explanation. As Evan, Janie, and Franny continue their effort to understand why some finches survive and others do not, they eventually verbalize the idea that differences in the finches' beaks could determine their survival. This idea runs through their group discussions from a very early point of their investigation, although they do not decide to record the hypothesis until the third day of their investigation. Having examined field notes in the GF environment that indeed show that different birds eat different seeds, they are unsure of the full implications of these data, but decide they are important to record. So, they decide to create a new explanation, and this requires them to choose a specific explanation guide. The need to choose an explanation guide ignites a debate about a fundamental evolutionary concept.

Evan: It's still an environmental catastrophe. Oh no wait, no its not. It's a—selective pressure. [reading list of available explanation guides]

Franny: Is it that one?

Evan: Environment causes—[reading description of "selective pressure" guide]

Janie: No!

Evan: Yeah, to be selected for

- Janie: Yeah, but that means like . . .
- Evan: //what food they eat//
- Janie: —organism with these trait
- Evan: //the trait being the food
- Franny: Yeah, that's right.
- Janie: No, because like, if my trait is to eat steak and there's no steak, I'm immediately gonna go to something else.
- Evan: If you're only a vegetarian and you only eat—you don't eat meat, you're not gonna eat meat. Well, that depends//
- Janie: Are you insane?!
- Franny: OK, OK. Don't think of people. Think of these guys. If they only eat one type of seed with their beaks and that seed is gone then they can't live anymore.

The need to choose a particular guide provokes the students to map their thinking about the problem into a disciplinary framework, and helps to make key domain concepts problematic. This exchange surfaces confusion about what constitutes a trait, and provides the group an opportunity to clarify the idea for themselves. Through their discussion the group converges, at least momentarily, on a common idea of what a trait is (cf. Roschelle, 1992).

This was a singular event for these students, the only time we observed them to debate the meaning of a particular explanation guide. In other situations, and in other groups, students seemed generally to decide without discussion which explanation guide they would use. One student would suggest a guide, and others in the group would silently assent. Negotiation then occurred in the context of what the group would actually write for an explanation. Figure 4 shows a version of the explanation that Franny, Evan, and Janie proposed for finch survival that masks the extended discussion the group had about what exactly they were claiming. Such negotiations occurred across all of the groups we observed, although again with a fair amount of variation in both frequency and duration. Negotiations seemed somewhat determined by whether or not the student in control of the keyboard was the dominant member of the group and how engaged other students were.

Evidence Evaluation. A fourth kind of epistemic discourse that occurred mostly within ExplanationConstructor includes the ways in which students evaluated their explanatory claims in terms of the evidence that they did or did not have. After selecting their new explanation's template, Franny, Janie, and Evan jointly constructed an initial explanation for why some finches survived (Figure 4). Evan reiterates their working hypothesis, that certain finches eat certain seeds, and these seeds may have been killed by the drought.

- Evan: Certain finches might eat certain seeds, and these plants might have been killed in the drought.
- Franny: [nodding her head in agreement]. Certain finches—well, are you saying because—because of their beak lengths, or because of what—?
- Evan: We don't know that yet.

Evan's assertion that they do not know the cause of the finches' survival reflects the current state of the group's investigation, that at this point they had not yet looked specifically at the

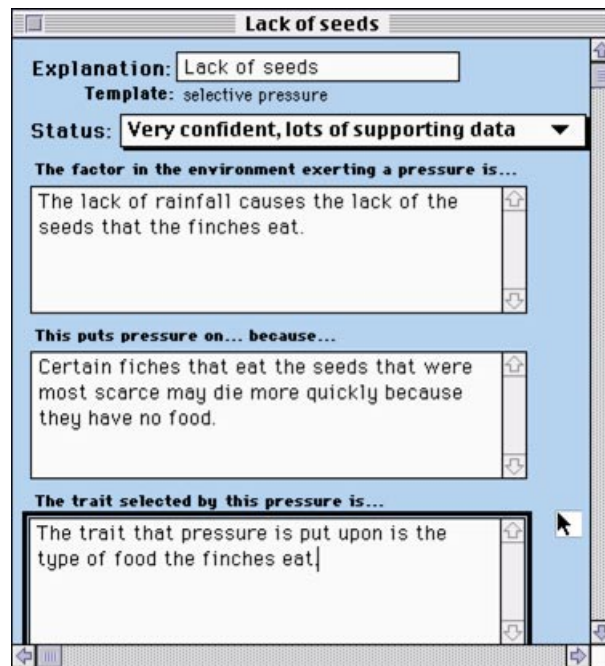


Figure 4. Janie, Franny, and Evan's initial explanation for why some finches can survive.

available data about beak lengths. They had, on the other hand, extensively explored field note data about individual finches' eating behaviors that had showed them that different birds were eating different seeds. Evan is thus holding the group to a strict standard of evidence, claiming that they do not know that the beak lengths are the causal factor that distinguishes "certain finches" from others. This issue of evidence arises as the students are negotiating their explanation.

Again, in other groups this sort of talk was common although its occurrence was highly variable. All of the groups we observed spent a lot of time talking about the data they were looking at. Across all of these groups there were at least a few instances in which students explicitly evaluated the status of their claims with respect to available data, as in this example. That such evaluations happened reflects the nature of the task and suggests that the demand to record developing explanations may have encouraged this consideration. What we rarely saw in these conversations, however, was explicit talk about how specific data could be linked to specific claims. Students talked a lot about how to interpret the data, but rarely how they could use data to support their explanations.

Recognizing Important Data. A final common practice among the groups we observed was the recognition that data that they were looking at were important to their ongoing investigation. Such moments were critical to the path of investigation, as they either supported a current idea or led students to a new idea. These conversations were nearly always initiated by students' examination of new data within an investigation environment. Yet, once data were recognized as important, they often, but not always, led students to return to ExplanationConstructor to note that importance. For example, with Evan, Franny, and Janie, once they noticed the lack of rainfall in the dry season of 1977 (Figure 2) they immediately decided to create their "Rainfall" explanation.

At first glance it may not seem that recognizing a piece of data is epistemic. We label it this way to emphasize features of the investigations students were carrying out here. First

of all, students generated dozens of pieces of data in each investigation, out of hundreds of possible choices. Consequently, students had to purposely select some data as important over others. We would argue that this is an epistemic decision, made in terms of students' current working explanation and their judgment of new data's impact on that. On average, student groups in these classes constructed less than two explanations for each problem, so clearly not every piece of data they looked at stood out as significant. One limit of our analyses here is that we focused more on where these decisions were being made, in order to clarify ExplanationConstructor's role in them, than on how.

Summary. We have presented Evan, Franny, and Janie as an example to provide a sense for how students' collaborative discourse through the course of an investigation was influenced by features of ExplanationConstructor. We cannot claim that these students could not possibly have had such epistemically focused conversations without using ExplanationConstructor. Their dialogue suggests that the explicit epistemic representations of the software shaped these conversations in valuable ways. We suggest this happened because these representations were (a) persistent and (b) epistemically grounded within the discipline. They highlighted the epistemic goals while articulating important domain-specific frameworks for explaining things. Because their questions and explanations were always available and open to inspection, students could focus their conversations on the issues of what they knew, and how they knew it. There are limits to these scaffolds, of course. Namely, they cannot help students to interpret specific data, nor can the software interpret students' ideas. Moreover, our observations suggest that students varied in their readiness to engage the tools as intended. We consider these limits and their implications in the discussion.

These case studies provide some initial evidence that ExplanationConstructor can provide strategic guidance during inquiry in ways we expected given our conceptual framework. They are not yet sufficient to explicate the variability in students' use of such guidance. Besides this variation in students' interactions with the tool, students' success explaining the problems varied (see Sandoval, 2003). This variability in the constructed artifacts and in students' investigations suggests that explanation evaluation has a crucial role to play both in helping students to understand the specific problems they investigate and the more general "game" of constructing scientific explanations. Our intended reflection activities were not enacted in this study as planned. First, Mr. Gray chose to skip the midpoint critiques because he was concerned about the time they might take and wanted to finish the unit quickly. The postassessment activities we designed did not seem to us very effective. Perhaps not surprisingly, students rated their own performances highly, and reflective prompts did not generate specific reflections on explanation quality. Following these assessments, pairs of groups asked each other to assess their explanations, and were given "interview" prompts to elicit specific reflections about their explanations. These paired reflections did not promote substantive reflection (Sandoval & Reiser, 1997). The ongoing reflection during explanation construction did not extend beyond students' work on the computer.

Year 2: Developing Support for Explanation Evaluation

Our findings from our first classroom trials encouraged us to revise ExplanationConstructor and our curriculum in several ways to foster more explicit explanation evaluation. First, we wanted students to cite more data in their explanations, to be more explicit about how their claims relied upon or explained specific data. Second, we wanted students to more actively reflect upon their own and their peers' performance in epistemic terms, by

considering whether or not claims were clearly articulated, made sense, held together, and were supported with data. Finally, we wanted such epistemic conversations to occur more publicly in the classroom, to extend beyond students' small group work.

To encourage more explicit data citation, we revised the earlier version of ExplanationConstructor (as seen in Figure 3) to the version shown in Figure 1. The key change here was to make potential evidence immediately visible (see right side of Figure 1) and enable data to be cited directly in the explanation texts that students wrote. We expected this would encourage more data citation, and more precise citation. This would make it easier to understand the data on which students justified their claims and why they thought such data mattered. We also added a reviewing feature to ExplanationConstructor to facilitate peer- and self-assessment activities described above (mid- and postinvestigation reviews). Also, in collaboration with Mr. Gray, we developed a written rubric of criteria for good explanations, as shown in Table 2. Another important change was a redoubled effort in the classroom to conduct both the mid- and postinvestigation critique and assessment activities.

Setting. In Year 2, we returned to work with Mr. Gray and a colleague he recruited to work with us. Mr. Gray was teaching only two honors biology classes, so we recruited Mr. Brown to examine our revisions in two regular level classrooms as well. There were 87 students in the four classes (43 boys, 44 girls). The purpose of the study was to examine two key questions. First, would the revisions we made to ExplanationConstructor and the curriculum encourage students to use data more explicitly in their explanations? Second, would

TABLE 2
Epistemic Components of Explanation Rubric Used in Year 2

Criterion	Description for Students
Thoroughness and clarity of explanations	Scientific explanations are causal explanations. They are stories about how one thing causes another thing. They explain how or why things happen. Most scientific explanations involve chains of cause and effect: A causes B which causes C which causes D. This part of your score will be based on how clearly you state the causal chain in your explanation.
Use of data	Scientific explanations are scientific because they are based on patterns of data. You will be graded on your rationale for how you link the data to support your explanations. Within the computer program you will be able to cite data and specifically link data (measurements, graphs, weather conditions, populations) to support parts of your explanations.
Ruling out alternative explanations	Like most scientists, your group is certain to reject ideas along the way to determining what you think is the best explanation. You cannot be sure you have the best explanation if you haven't considered alternative explanations and documented why those explanations should be rejected in favor of your best explanation. Your group will write at least two articulated explanations.
Documenting the limitations of your explanations	Any explanation, no matter how thorough it seems, will not be able to account for all available data or be weakened by missing data. Your group will be responsible for documenting the limitations of each of your explanations. If there are limitations that are undocumented, you will lose points.

the revised supports for reflection promote students' engagement in epistemic practices of explanation evaluation? We again found that students were quite able to use ExplanationConstructor to construct explanations of natural selection, for both the GF and TB problems. In both problems, students argued for specific traits providing advantages. They also used data to support their claims more than in the previous year (Sandoval, 2001).

Epistemic Practices of Evaluation. To understand students' practices of explanation evaluation in the revised learning environment, we collected the peer critiques and self-assessments that students conducted during and after both the GF and TB investigations. Students used the evaluation rubric (Table 2) to guide both their critiques of other groups' explanations and self-assessments of their own, although both teachers focused self-assessments on documenting limitations to explanations. We analyzed the texts of both peer critiques and self-assessments to identify the types of critiques and limitations students offered of their own and each others' work.

Peer Critiques. Each group critiqued one other group in their class at roughly the midpoint of both the GF and TB investigations. Groups simply rotated to the nearest computer to review another group. Groups did not talk to other groups, but discussed within their group each explanation. They were instructed to use the rubric handed out in class, and to be constructively critical. Groups were told not to tell the other group what they thought the answer was, but to indicate to the group being reviewed what parts of their explanation needed to be more clear, needed more detail, needed evidence, and so on. Students wrote their critiques in a text box provided by ExplanationConstructor (lower left in Figure 1). Individual statements were transcribed and coded for the type of critique. Codes for types of critiques emerged from these data, although they are similar to other epistemically oriented analyses (e.g., Jiménez-Aleixandre, Rodríguez, & Duschl, 2000; Kelly, Druker, & Chen, 1998b; Resnick et al., 1993).

Table 3 shows the types of critiques we observed in this sample, across both GF and TB problems. The most common critiques were that groups failed to provide enough data,

TABLE 3
Peer Critiques^a of Explanations for Both Problems, All Classes, in Year 2

Critique	Definition	N (%)
Lack of data	Point out a lack of data to support a claim; including no data	26 (23.85)
Lack of mechanism	Lack of explanation for how something happens	22 (20.18)
Affirmation	Compliment or other positive remark	15 (13.76)
Counter claim	Offer a counter claim for interpretation of data	11 (10.09)
Lack of information	Explanation lacks information, unspecified	11 (10.09)
Say more	A claim that more needs to be said, or more work needs to be done	9 (8.26)
Alternative claim	Offer an alternative that needs to be considered	7 (6.42)
Objection	A generalized objection to a claim	3 (2.75)
Suggestion	A suggestion about investigative method, or rhetorical organization	3 (2.75)
Need for warrant	A vague claim for the need to "show" something	2 (1.83)

^aOrdered by decreasing frequency of occurrence.

including citing no data at all, or that groups failed to adequately explain the mechanism behind a claim. These latter critiques could be fairly general, for example, “they don’t explain why this happened.” Alternatively, they could be specific, such as, “how do you think that the mutation affected the TB strain . . . ?” Remember that these critiques occurred in the middle of students’ work, so a lack of evidence for claims as well as these other noted deficiencies reflect the fact that groups were not done and differed in the amount of progress they had made. Also relatively common were counter claims, where students asserted that some other interpretation of proffered data was possible, and general claims that explanations lacked sufficient information. All of these critiques suggest that students were able to apply the criteria in their rubric. These four types account for more than 60% of all critiques. Somewhat surprisingly, a common critique included simply an affirmation that a group had done a good job. Sometimes these affirmations were specific, as in “good job explaining your data,” and other times they were simple compliments, “Wow . . . good job!” Overall, students critiques suggest they could interpret and apply the evaluation rubric they had been given.

Self-Assessments. Following each investigation, groups assessed the limitations of their own explanations. Mr. Gray and Mr. Brown organized these activities similarly, having students record limitations to their final explanations either within their ExplanationConstructor journals or on separate worksheets. Stated limitations were transcribed and coded according to the nature of the critique. Again, our approach was to let codes emerge from students’ statements. We found a number of limitations that were rather different from the earlier peer critiques. Table 4 shows the types of self-assessed limitations and their frequency in our sample. The majority of the assessments students made of their own work were claims of the limitations to what they were able to find out. On the surface these assessments appeared quite similar to the kinds of limitations scientists might level against their own work, but the sources were quite different.

The most common critique raised by students, nearly a third of all assessed limitations, were that students were limited in what they were able to find out because the computer

TABLE 4
Self-Assessed Limitations^a of Explanations for Both Problems, All Classes, from Year 2

Limitation	Definition	N (%)
Verisimilitude	Attribute lack of certainty to data missing from computer	34 (30.91)
Alternative claim	Suggest alternative claim, other than own. Sometimes spurious	28 (25.45)
Limited extent	Acknowledge limitation to extent of knowledge about problem	17 (15.45)
Restatement	Reassert their own claim and/or evidence for it	17 (15.45)
Ambiguity in data	Acknowledge uncertainty from ambiguity in data	6 (5.45)
Question	Raise an open question	3 (2.73)
Assumption	Question an assumption underlying own explanation	2 (1.82)
Lack of time	Ran out of time to find evidence or research a claim	2 (1.82)
Absurd alternative	Make absurd alternative claim	1 (0.91)

^aOrdered by decreasing frequency.

environments lacked data or kinds of tests, i.e. verisimilitude to the real world. For example, it was common for students to mention that in the GF problem they were unable to rule out disease as the cause of the finches' deaths, because "we couldn't find any info." The critique is leveled against the computer environment rather than focused on the content of students' own work. We cannot say whether in such cases students seriously entertained a disease hypothesis and were frustrated by not being able to explore it, or if they thought that in the real world they would be able to get all possible data about all possible causes, or if they were just going through the motions of proposing and refuting alternative explanations as they had been asked.

A second common claimed limitation was to suggest a possible alternative to students' own answer. We distinguished these limitations from verisimilitude if the statement made no mention of data or potential data (e.g., "there could have been a disease that killed the finches"). Considering such alternatives was part of the rubric students were working under, although many of these claims seemed to us to be invented simply to produce an alternative. These two kinds of claimed limitations make up over half of all the assessed limitations in the sample, and both suggest a strategy of locating any limitations on sources outside the students' own explanations. This was not what we had intended, but is perhaps not unexpected.

The next most common limitation was an acknowledgment of the limits to what students had learned (limited extent). The sources of this limitation were also generally tied to data students had not looked at, but they were framed in terms of students' actual work. That is, they were assessments of the open issues remaining in each problem. For example, in the TB problem students recognized that although they had discovered genetic differences between strains, they were not able to learn how these differences changed molecular structures. Although they could not explore such structures in the software, such a statement locates the limitation in the explanation, not the software. Other assessed limitations that also examined the status of students' own explanations included acknowledgments of ambiguity in some data, raising questions that they had been unable to answer, and noting that their own explanations rested on certain assumptions. As can be seen in Table 4, these latter assessments were rare in this sample.

Overall, students' self-assessed limitations seemed primarily aimed at justifying their own work, by placing limitations outside of their explanations, rather than frankly noting what they did and did not know from their investigations. One likely reason for this, given the typical context of school science, is that students saw limitations as an opportunity to either reaffirm that they had come to the "right" answer for the problem, or to suggest why forces beyond their control hindered their solution. Admittedly, these data are insufficient to conclude this was the case. We can say that in this context, given the opportunity to assess the limitations of their own explanations, students were more likely to look outside their explanations rather than reflect directly on their own claims and evidence.

Summary. Students' self-assessed limitations contrast with their critiques of each other's work, which were more directly focused on the content of peers' explanations. Consequently, the peer critiques were more likely to focus on what we value as epistemic practices. It is possible that having students assess their peers' explanations rather than their own might have encouraged that focus in the postinvestigation reviews. At the same time, these claimed limitations suggest that students understand the general rules of the game in evaluating these explanations. Critiques of the computer environments' data sets and recognition that some claims could not be refuted each reflect an understanding that data are needed, or are at least preferable, for supporting claims. In Year 2, by focusing more directly on evaluative reviews, and providing students with some guidance about the nature of such reviews, we were

successful in getting students to evaluate their written explanations in epistemic terms. We also saw that differences in task structure affected how students applied epistemic criteria.

DISCUSSION

The trajectory of our research and design program has been to develop an approach for building learning environments that can support students' inquiry and their understanding of the nature of science. We have chosen here to focus on the conceptual framework underlying our design and to present some of the data that have helped us to refine our design and our understanding of the learning issues at hand. Our goal in these studies has not been to evaluate the effectiveness of ExplanationConstructor or our curricula. If that were the goal, our methodologies would have been poorly suited to the task. Instead, our goal has been to understand whether or not our tools appear to function in their intended settings in ways predicted by the conceptual framework. Here, our methodological approach so far has been a starting point, a way to refine our analytic attention.

Understanding how specific technological tools function in complex settings such as classrooms is tricky business. Because the intervention is complex and the existing setting is complex, analysis has to try to account for the entire package rather than isolate the effects of particular pieces—the pieces have no effects in isolation (Salomon, 1996). This begs the question, however, of how we can tell whether or not the software is worth refining, or how to recreate successful aspects of the environment. Our answer to the first question comes from looking at students' behavior within the learning environment and determining how a specific tool, like ExplanationConstructor, is implicated in that behavior. Our answer to the second part of the question remains a focus of further research. Still, our findings suggest some of the roles that a tool like ExplanationConstructor can play in supporting inquiry, and some of the limits to technological support. ExplanationConstructor is an example of an epistemic tool, one that can help students articulate their thinking in ways meaningful within disciplines. Our studies so far suggest the roles that epistemic tools can play within science inquiry learning environments, and how such tools can contribute to the study of students' scientific epistemologies.

Roles of Epistemic Tools

ExplanationConstructor is just one of several recently developed epistemic tools that structure students' efforts to construct arguments (Bell & Linn, 2000; Scardamalia & Bereiter, 1993–1994; Toth, Suthers, & Lesgold, 2002) or models (Jackson et al., 1994). As opposed to what we might call conceptual tools, such as visualizations and simulations designed to help students reason about specific phenomena, epistemic tools help students articulate their understanding about such things in ways that are congruent with selected epistemological commitments in science. Epistemic and conceptual tools can work together to support students' inquiry. In our own work, for example, the specific investigation environments are mainly designed to help students understand concepts of natural selection, while ExplanationConstructor is primarily designed to help them articulate that understanding in a certain epistemic form (Collins & Ferguson, 1993). There are examples of conceptual tools embedded within curricula that pay close attention to epistemic aspects of inquiry (e.g., White & Frederiksen, 1998). We suggest, however, that epistemic tools have unique roles to play in supporting a kind of inquiry that both develops students' ability to engage in scientific practice and their epistemological conceptions of that practice.

Epistemically Structured Articulation of Student Thinking. One of the ways that epistemic tools can foster scientific practice is by structuring students' articulation of their thinking in epistemically valued ways. This can help students to attend to important epistemic goals during inquiry. For example, the high level hierarchy of the ExplanationConstructor journal organizes students' work in terms of questions and their explanations. As seen in the case study above, this organization supported students' focus on the goals of their inquiry. In turn, this helped the group to set and pursue investigative strategies in relation to their explanatory goals. The group's monitoring of their progress was not simply metacognitive, but was conducted in epistemic terms. Had they answered the current question? Had they finished the current explanation? What evidence did they have; did they need? Such questions cannot be answered in any sort of general way, they depend on some understanding of what counts as an explanation, and as evidence, within the particular domain. Asking and answering such monitoring questions was encouraged by the framing of students' inquiry, as we discuss below, and was supported by the software. These students repeatedly referred to the questions they had articulated, the explanation prompts in the guides they chose for their explanations, and negotiated key disciplinary ideas when writing their explanations. This epistemic cognition may be especially important for solving ill-structured problems (Kitchener, 1983), as authentic inquiry problems commonly are. The students we observed exploited the fact that their questions and emergent explanations were persistent, always there and always accessible.

These explicit epistemic supports also structured the form of students' explanations in ways that supported subsequent evaluation. During both critique and assessment activities, the software helped to make clear when claims were supported with evidence or not, which questions were answered, and so on. As we learned in Year 2, such software scaffolds are better used when embedded within a clear task structure with other needed material supports (e.g., the rubric of Table 2). The importance of such evaluation is that it makes epistemic criteria explicit and a public part of classroom discourse. Explanation construction itself seems to reflect students' epistemological ideas, but not necessarily provoke change in them (Sandoval, 2003). Instead, debate about which explanations are better than others and why seems to be a key aspect of developing students' epistemic criteria (Rosebery, Warren, & Conant, 1992).

Tool Use Framed in Knowledge-Building Activities. We cannot overemphasize the importance of the fact that students' use of ExplanationConstructor and other tools in this curriculum was embedded within tasks that were specifically aligned with the epistemic practices we are trying to develop. This point may seem obvious, but we stress it because it is crucial to understand that students' performances here do not originate from the tools but are supported by them. Vygotsky (1978) argued that tools mediate human thought and activity. We intentionally designed ExplanationConstructor to mediate students' inquiry activity in particular ways. Students' use of the software in intended ways depended upon their understanding of the purpose of the tool and its affordances for action. This understanding was communicated in part by the teachers and their framing of students' inquiry.

The teachers in these classrooms set the expectations for student performance, both in this particular unit and more generally throughout the course. Students interpreted their understanding of what their teachers wanted during their specific investigations here in terms of their understanding of their teachers' expectations for previous assignments. That students successfully articulated causal explanations and cited data to support their claims was partially a result of their teachers' staking those as the performance expectations. This was especially true in Year 2, when we and the teachers reified these expectations in an

explicit rubric for students. One of the consequences was that students cited much more data than their peers from the previous year (Sandoval, 2001). We also believe that changes to the software that made the relation of evidence to claims more visible, and more salient, contributed to the greater citation of data. The tool acts as an enabler and support to activity whose value comes from established classroom norms.

There has been a repeated call among science educators in recent years to make argumentation a central practice of instruction (Driver et al., 2000; Duschl, 1990; Kuhn, 1993). Our own effort to do this suggests the potential for software environments to support argumentation, as well as their limits. In our classrooms, we find that epistemic discourse is largely limited to small group settings where technological or material supports encourage it. Sustaining an explicitly epistemic discourse over time seems to require a shift toward scientific argumentation as the public norm of the classroom (Khishfe & Abd-El-Khalick, 2002; Rosebery, Warren, & Conant, 1992; Smith et al., 2000). In these more public contexts, the artifacts produced with epistemic tools can play an instrumental role in grounding discourse, but the major form of support needed in these contexts seems largely social. Further research is needed to understand how these socially supported norms emerge and can be sustained (cf. Cobb et al., 2001; Tabak & Reiser, 1997).

Epistemic Practices and Epistemological Beliefs

We have defined epistemic practices as the reasoning and discursive practices involved in making and evaluating knowledge, in this case scientific knowledge. From this effort to support students' construction of scientific knowledge, what can we say about the epistemic practices students engage in? What do these practices say about students' underlying ideas about the nature of scientific practice? Through our experiences we have come to understand that an important research goal arising out of our design work is the opportunity it provides to study school science-in-the-making (Kelly, Chen, & Crawford, 1998a), an aspect of epistemological studies increasingly recognized as needed (Kelly, Chen, & Crawford, 1998a; Kelly & Duschl, 2002; Lederman, Wade, & Bell, 1998).

The conversations we see students have as they collaboratively investigate specific problems is attentive to epistemic concerns. They question how they know what they know, and whether they have answered their driving questions. As Janie, Franny, and Evan illustrate, students evaluate their progress both in terms of whether or not they have articulated causal claims and whether or not they have evidence to support such claims. Given that students do not often spontaneously evaluate their progress in these ways (Krajcik et al., 1998; Kuhn, Schauble, & Garcia-Mila, 1992), the level of support provided by our learning environment (tools, materials, and activity structures together) may be responsible. We point out that this remains to be more systematically investigated. We can say that in the classrooms where we have intervened, students have demonstrated certain fundamental practices of science. Their conversations as illustrated here, and our analyses of their explanations (Sandoval, 2001, 2003), show that students work to articulate causal accounts, strive to make such accounts cohere, and argue that claims ought to be based upon evidence.

In the substantial literature on students' understanding of the nature of science over the last 50 years, a consistently dismal picture of students' epistemological beliefs has been drawn. From surveys and interviews that probe students' ideas about professional (or formal) science, students rarely see theories as creative ideas that are developed to explain observed data. Instead, they see data as definitively able to determine if an idea is right or wrong, or sometimes say that data *is* the answer to a question (Carey et al., 1989; Driver et al., 1996; Lederman, 1992). This work has largely been guided by an assumption that students' epistemological beliefs are stable, consistent frameworks independent of particular subject

areas (e.g., Smith et al., 2000). Recently, however, evidence has been accumulating that suggests that students' conceptions of the nature of science may not be stable, coherent belief systems. Instead, they may be self-contradictory (Roth & Roychoudhury, 1994; Sandoval & Morrison, 2003), inconsistent across contexts (Leach et al., 2000), and conflated with personal conceptions of the best ways to learn science (Hammer, 1994; Hogan, 2000). This growing evidence that students' epistemological conceptions of science may be fragmented has led to the development of a view of epistemological knowledge as localized "resources" (Hammer & Elby, 2001) used to reason in particular situations.

Abd-El-Khalick, Lederman, and their colleagues have recently argued that epistemological development must be an explicit instructional goal (Abd-El-Khalick & Lederman, 2000; Lederman et al., 2002) and that explicit instruction can change students' conceptions (Khishfe & Abd-El-Khalick, 2002). We agree, although an outstanding issue remains connecting students' practices in school to their expressed beliefs of nature of science. The sort of environments we have constructed here provide rich contexts in which to study students' practice, as they encourage students to engage in science in a way that is more like authentic practice than typical instruction. Beyond documenting students' practices in such environments, future research must also explore the conceptions of science that students have that motivate their performance. Rather than asking them abstract questions about a professional science they have no experience with, the artifacts that students construct during their own inquiry may provide a more accessible context through which to assess underlying epistemological ideas. Doing so will lead both to better theoretical accounts of epistemological development as well as better instructional practices for inquiry.

We are deeply grateful to the teachers and students who participated in these studies. We thank the BGuILE team for their contributions to this work, especially Iris Tabak, Richard Leider, and T.J. Leone. Thanks also to Yasmin Kafai, Mike Rose, Nancy Brickhouse, and three anonymous reviewers for their helpful comments on earlier drafts.

REFERENCES

- Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education*, 22(7), 665–701.
- 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.
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27, 415–427.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, 2(2), 141–178.
- Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68(4), 493–503.
- Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). 'An experiment is when you try it and see if it works': A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, 11, 514–529.
- Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, 28(3), 235–251.
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences*, 10(1/2), 113–163.
- Collins, A. (1992). Toward a design science of education. In E. Scanlon, & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). Berlin: Springer-Verlag.

- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Erlbaum.
- Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist*, 28(1), 25–42.
- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8.
- Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham, UK: Open University Press.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84, 287–312.
- Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, 17, 397–434.
- Duschl, R. A. (1990). *Restructuring science education: The importance of theories and their development*. New York: Teachers College Press.
- Erickson, F. (1992). Ethnographic microanalysis of interaction. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 201–225). New York: Academic Press.
- Gitomer, D. H., & Duschl, R. A. (1995). Moving toward a portfolio culture in science education. In S. M. Glynn, & R. Duit (Eds.), *Learning science in the schools: Research reforming practice* (pp. 299–326). Hillsdale, NJ: Erlbaum.
- Greene, E. D. (1990). The logic of university students' misunderstanding of natural selection. *Journal of Research in Science Teaching*, 27(9), 875–885.
- Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12(2), 151–183.
- Hammer, D., & Elby, A. (2001). On the form of a personal epistemology. In P. R. Pintrich (Ed.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Erlbaum.
- Hogan, K. (2000). Exploring a process view of students' knowledge about the nature of science. *Science Education*, 84(1), 51–70.
- Jackson, S. L., Stratford, S. J., Krajcik, J., & Soloway, E. (1994). Making dynamic modeling accessible to precollege science students. *Interactive Learning Environments*, 4(3), 233–257.
- Jiménez-Aleixandre, M. P., Rodríguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, 84, 757–792.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39–103.
- Kelly, G. J., Chen, C., & Crawford, T. (1998a). Methodological considerations for studying science-in-the-making in educational settings. *Research in Science Education*, 28(1), 23–49.
- Kelly, G. J., Druker, S., & Chen, C. (1998b). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. *International Journal of Science Education*, 20(7), 849–871.
- Kelly, G. J., & Duschl, R. A. (2002). Toward a research agenda for epistemological studies in science education. Paper presented at the Annual Meeting of NARST 2002, New Orleans, LA.
- Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching*, 39(7), 551–578.
- Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition. *Human Development*, 26(4), 222–232.
- 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. *Journal of the Learning Sciences*, 7(3/4), 313–350.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319–337.

- Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. San Diego, CA: Academic Press.
- Kuhn, D., Schauble, L., & Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. *Cognition and Instruction*, 9(4), 285–327.
- Lajoie, S. P. (1993). Computer environments as cognitive tools for enhancing learning. In S. P. Lajoie, & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 261–288). Hillsdale, NJ: Erlbaum.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Leach, J., Millar, R., Ryder, J., & Séré, M.-G. (2000). Epistemological understanding in science learning: The consistency of representations across contexts. *Learning and Instruction*, 10, 497–527.
- Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29(4), 331–359.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of Nature of Science Questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497–521.
- Lederman, N. G., Wade, P. D., & Bell, R. L. (1998). Assessing the nature of science: What is the nature of our assessments? *Science & Education*, 7, 595–615.
- Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex.
- Linn, M. C., & Songer, N. B. (1993). How do students make sense of science? *Merrill-Palmer Quarterly*, 39(1), 47–73.
- Mayr, E. (1988). *Toward a new philosophy of biology*. Cambridge, MA: Harvard University Press.
- Norman, D. A. (1988). *Psychology of everyday things*. New York: Basic Books.
- NRC. (1996). *National science education standards*. Washington, DC: National Academy Press.
- O'Neill, D. K., & Gomez, L. (1994). *The Collaboratory Notebook: A distributed knowledge-building environment for project-enhanced learning*. Paper presented at the Ed-Media 94, Charlottesville, VA.
- Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47–87). Cambridge, MA: Cambridge University Press.
- Reiser, B. J. (2002). Why scaffolding should sometimes make tasks more difficult for learners. Paper presented at the CSCL2002. Computer Support for Collaborative Learning, Boulder, CO.
- Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver, & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Erlbaum.
- Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction*, 11(3/4), 347–364.
- Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences*, 2(3), 235–276.
- Rosebery, A. S., Warren, B., & Conant, F. R. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *Journal of the Learning Sciences*, 2(1), 61–94.
- Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research*, 68(1), 35–59.
- Roth, W.-M., & Roychoudhury, A. (1994). Physics students' epistemologies and views about knowing and learning. *Journal of Research in Science Teaching*, 31(1), 5–30.
- Rutherford, F. J., & Ahlgren, A. (1990). *Science for all Americans*. Washington, DC: AAAS.
- Salomon, G. (1996). Studying novel learning environments as patterns of change. In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology-supported learning environments* (pp. 363–377). Mahwah, NJ: Erlbaum.
- Sandoval, W. A. (2001, April). Students' uses of data as evidence in scientific explanations. Paper presented at the Annual Meeting of the American Educational Research Assn, Seattle, WA.

- Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, 12(1), 5–51.
- Sandoval, W. A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching*, 40(4), 369–392.
- Sandoval, W. A., & Reiser, B. J. (1997, March). Evolving explanations in high school biology. Paper presented at the Annual Meeting of the American Educational Research Assn, Chicago, IL.
- Scardamalia, M., & Bereiter, C. (1993–1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, Special Issue: Computer Support for Collaborative Learning, 3(3), 265–283.
- Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences*, 4(2), 131–166.
- Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences*, 1(2), 201–238.
- Settlage, J. (1994). Conceptions of natural selection: A snapshot of the sense-making process. *Journal of Research in Science Teaching*, 31(5), 449–457.
- Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition & Instruction*, 18(3), 349–422.
- Tabak, I., & Reiser, B. J. (1997). Complementary roles of software-based scaffolding and teacher-student interactions in inquiry learning. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Proceedings of Computer Support for Collaborative Learning'97* (pp. 289–298). Mahwah, NJ: Erlbaum.
- Tabak, I., Smith, B. K., Sandoval, W. A., & Reiser, B. J. (1996). Combining general and domain-specific support for biological inquiry. Paper presented at the *Intelligent Tutoring Systems'96*, Montreal, Canada.
- Tobin, K., Tippins, D. J., & Hook, K. S. (1995). Students' beliefs about epistemology, science, and classroom learning: A question of fit. In S. M. Glynn, & R. Duit (Eds.), *Learning science in the schools: Research reforming practice* (pp. 85–110). Mahwah, NJ: Erlbaum.
- Toth, E. E., Suthers, D. D., & Lesgold, A. (2002). "Mapping to know": The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education*, 86, 264–286.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition & Instruction*, 16(1), 3–118.
- Windschitl, M., & Andre, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching*, 35(2), 145–160.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, 20, 99–149.