COGNITION AND INSTRUCTION, *18*(3), 349–422

Sixth-Grade Students' Epistemologies of Science: The Impact of School Science Experiences on Epistemological Development

Carol L. Smith
Department of Psychology
University of Massachusetts at Boston

Deborah Maclin and Carolyn Houghton
Educational Technology Center
Harvard Graduate School of Education

M. Gertrude Hennessey
Science Department
St. Ann School

Previous studies have documented that middle school students have a limited "knowledge unproblematic" epistemology of science (i.e., scientists steadily amass more facts about the world by doing experiments) with no appreciation of the role played by scientists' ideas in guiding inquiry. An important question concerns to what extent students this age and younger are ready to restructure their epistemological views to focus on more "constructivist" issues: the conjectural, explanatory, testable, and revisable nature of theories. This study tests the claim that even elementary school students can make significant progress in developing a more sophisticated, constructivist epistemology of science, given a sustained elementary school science curriculum that is designed to support students' thinking about epistemological issues. To assess the impact of elementary science experiences on students' epistemological views, 2 demographically similar groups of 6th-grade students were individually interviewed using the Nature of Science Interview developed by Carey

Requests for reprints should be sent to Carol L. Smith, Department of Psychology, University of Massachusetts at Boston, 100 Morrissey Boulevard, Boston, MA 02125. E-mail: carol.smith@umb.edu

and colleagues (Carey, 1991; Carey, Evans, Honda, Jay, & Unger, 1989). Both groups had experienced sustained elementary science instruction; 1 taught from a constructivist perspective and 1 taught from a more traditional perspective. We found that students in the more traditional science classroom had developed a knowledge unproblematic epistemology of the type previously reported by Carey et al. (1989). In contrast, students in the constructivist classroom had developed an epistemological stance toward science that focused on the central role of ideas in the knowledge acquisition process and on the kinds of mental, social, and experimental work involved in understanding, developing, testing, and revising these ideas. We conclude that elementary schoolchildren are more ready to formulate sophisticated epistemological views than many have thought. We discuss how these findings relate to the broader epistemological literature, and the features of the constructivist classroom environment that may have supported the development of these sophisticated understandings.

This study tests the claims that elementary school students (a) have coherent epistemological commitments, and (b) can make significant progress in developing a sophisticated, constructivist epistemology of science when taught science using a constructivist pedagogy. By *epistemology of science,* we are referring to the network of ideas students have about how knowledge is acquired and justified in science. By a *sophisticated, constructivist epistemology,* we mean an epistemology in which students are aware of the central role of ideas in the knowledge acquisition process and of how ideas are developed and revised through a process of conjecture, argument, and test. By a *constructivist pedagogy,* we mean a pedagogy in which students actively develop, test, and revise their ideas about how things work through collaborative firsthand inquiry with their peers. Students also reflect on their inquiry, and their inquiry and discussions are guided and scaffolded by a knowledgeable teacher.

One important goal of the K–12 science curriculum, currently advocated both by the National Research Council (1996) and the American Association for the Advancement of Science (AAAS, 1993), is for students to develop a sophisticated understanding of how knowledge is justified in science. This goal is important for several reasons. More sophisticated epistemologies may contribute to better learning of science content (Hammer, 1994; Schommer, 1993; Songer & Linn, 1991) and greater mastery of skills of argument (Honda, 1996; Kuhn, 1991; Sodian & Schrempp, 1997). More sophisticated epistemologies may also contribute to the development of informed citizens who understand the importance of reasoned argument in evaluating competing knowledge claims and who understand that the existence of genuine controversies in science does not undermine the value of scientific process and knowledge (Schwab, 1962). Yet, data from these and other sources (e.g., Driver, Leach, Millar, & Scott, 1996) indicate that many students do not achieve such sophisticated epistemologies.

DIFFICULTIES IN ACHIEVING THESE GOALS AMONG ELEMENTARY AND MIDDLE SCHOOL STUDENTS AND POSSIBLE REASONS FOR THESE DIFFICULTIES

Studies reviewed by Carey and Smith (1993) suggest that elementary and middle school students have very limited, fact-based epistemologies of science in which they make no explicit distinction between theory and evidence, let alone understand the interplay between theories and evidence in complex cycles of hypothesis testing.

For example, in the extensive scientific reasoning studies of Kuhn and colleagues (Kuhn, Amsel, & O'Loughlin, 1988), third-grade and sixth-grade students gave evidence of confusing theory and evidence in a variety of ways. When asked to provide evidence for their proposed theories, they often provided a further statement of their theory rather than referring to specific evidence. When evaluating their theories in light of specific evidence, they often ignored or overlooked evidence that was disconfirming to their theories and generally had to change their theories before they were willing to acknowledge the discrepant evidence. In reconstructing their proposed theories and observed evidence, they tended to bring the two into alignment, rather than acknowledge the discrepancies that had occurred. Finally, they were generally unable to state what pattern of evidence would disconfirm their theories.

More direct evidence that middle school students have limited epistemological views comes from interview studies probing their conceptions of science (Carey, Evans, Honda, Jay, & Unger, 1989) and scientific models (Grosslight, Unger, Jay, & Smith, 1991). For example, when asked a series of general questions about how scientists acquire their knowledge, seventh-grade students talked simply of doing tests, finding cures, and making observations, with no discussion of the interplay of ideas and evidence in the process. When asked about the nature and purpose of scientific models, students described a model as a little replica of some concrete thing, not as an abstract idea that guides the hypothesis-testing process. Furthermore, short-term teaching studies that have directly tried to teach these points (Carey et al., 1989; Carey & Smith, 1993; Honda, 1994) have produced, at best, modest changes in students' views.

Why do middle school students have such a limited understanding of the difference between theory and evidence in science? Why is it so hard for them to appreciate how theories guide the hypothesis-testing process even after participating in curricula that directly address these issues? In answering these questions, three kinds of explanatory factors need to be distinguished and considered. Students' difficulties may stem from (a) limitations in their prior schooling experiences in science; (b) a conceptual constraint imposed by their everyday epistemological views; or (c) a more biologically based, general developmental constraint on their thinking and reasoning. Let us consider the arguments for each kind of factor in turn.

On the one hand, science educators and researchers (Carey et al., 1989; Chinn & Malhotra, in press; Driver et al., 1996; Hodson, 1988; Metz, 1995; Nadeau & Desautels, 1984) have noted the limited epistemology informing most science educational practice as well as the limited opportunities students have to engage in theory-guided firsthand inquiry. They question practices that ask students to carry out recipe-like labs or activities without an understanding of the purpose of those activities and that ask students to memorize facts without an understanding of the framework informing those facts. Such educational practices tend to reinforce or consolidate a limited fact-based epistemology rather than help students to transcend it.

On the other hand, there is a burgeoning literature on students' general epistemological development (see Chandler, Boyes, & Ball, 1990; Hofer & Pintrich, 1997; King & Kitchener, 1994, for reviews) in which it has been argued that students start out with, at best, a limited, fact-based epistemology for thinking about knowledge claims in their everyday life. If these views function as a framework theory, then they should constrain the development of students' epistemological thinking about science and other domains.

Most of this research on epistemological development has been done with high school and college students. Students are typically presented with contrasting views on some controversial issue. Then they are asked to formulate and defend their own point of view, consider the status of the alternative point of view, and explain whether and how each controversy might be resolved. In all studies, there is broad agreement about the nature of students' starting epistemology (Hofer & Pintrich, 1997). That is, students begin by assuming knowledge is both simple (fact based) and certain. They do not justify beliefs based on an analysis of relevant evidence or detailed argument; rather, they simply assert their beliefs or give an authority-based justification. They do not acknowledge that differences of opinion may stem from different perspectives or frameworks; rather, they assume that these differences stem from inadequate knowledge, deception, or deceit, and will ultimately be resolved when all the facts are known or when one looks at the facts in an unbiased manner. Given that many high school students and noncollege-educated adults express such views, researchers have assumed that elementary school students would espouse similar or even more simplistic views (Hofer & Pintrich, 1997). One study that used a similar methodology with sixth-grade children confirmed that they see controversies as stemming from differences in specific knowledge rather than differences in interpretive framework (Kuhn et al., 1988).

Although we still know little about how to characterize elementary schoolchildren's epistemological views, the literature just mentioned suggests that they rest on ideas that are incommensurable with a sophisticated epistemology of science. If so, mastering the Nature of Science standards proposed by the AAAS (1993) with its notions of tentative, revisable, yet tested theories, would require students to restructure their epistemological concepts and to make fundamental conceptual changes.

Would elementary schoolchildren be able to make these conceptual changes if they were given extensive experiences in pursuing firsthand inquiry in science and negotiating the meaning of their findings among a community of learners? If they were able to make conceptual changes, what would their epistemology of science be like? Would they be able to develop a sophisticated, constructivist epistemology in which they appreciate that scientific knowledge is constructed through a process of conjecture, argument, and test? Prior to this study, there has been no way to answer that final question because the issues have not been directly investigated.

Some influential theorists of epistemological development have assumed that elementary students are incapable of such understandings because they have placed their work in Piagetian (Chandler, 1987) or neo-Piagetian (King & Kitchener, 1994) developmental frameworks, which assume the existence of a biologically based, general developmental constraint on students' thinking and reasoning. According to this view, elementary schoolchildren are "concrete" thinkers (Chandler, 1987; Inhelder & Piaget, 1958). Although they are capable of engaging in experimentation and learning from the observed results, they are incapable of reasoning hypothetically, understanding a theory as a conjecture involving unseen entities, examining the consistency of theoretical propositions, or deriving testable implications from such hypothetical conjectures. Children's initial absolutist epistemologies are thus seen as reflecting the limitations of concrete operational thought. Movement away from a fact-based and absolutist epistemology is an inherently late development, dependent on the achievement of formal operational thought, and often requiring the challenging intellectual experiences provided by college and graduate school. With the advent of formal operations, students become capable of more complex forms of perspective taking and reasoning. They are able to reflect on sets of beliefs of self and other, to identify these sets of beliefs as perspectives, and to consider how these perspectives influence one's interpretation of experience. These new abilities undermine students' belief in absolute truth.[1] They lead first to radical relativism, an epistemology in which all controversies are seen as reflecting legitimate differences in perspective, each supported by pieces of evidence, but with no means of resolving those differences. A more sophisticated constructivist view follows in which individuals are seen as "active constructors of meaning, able to make judgments and commitments in a relativistic context" (Hofer & Pintrich, 1997, p. 121). In King and Kitchener's (1994) ex-

[1]There is a disagreement between Chandler (1987) and King and Kitchener (1994) about the timing of these shifts and the exact factors responsible for them. Chandler argued that the shift to radical relativist views occurs during adolescence as a consequence of the development of formal operational thought. King and Kitchener argued that the shift generally occurs later during the college years. Although they believed that formal operational abilities are necessary for this development, they also believed that other kinds of educational experiences are needed to encourage the examination of epistemological issues and assumptions.

tensive studies, beliefs in absolute truth are typical for high school students, radical relativist views are common during the college years, and more sophisticated constructivist views are typical only among advanced graduate students, especially in the social sciences.

However, it is also possible that significant growth in epistemological thought can occur during the elementary school years. Work in developmental psychology over the last 20 years has shown that preschool and elementary schoolchildren have much greater logical and intellectual sophistication than Piaget and other developmental theorists had assumed, calling into question the whole construct of concrete and formal operational thought (see Carey, 1985a, and Metz, 1995, for reviews). This work has established that even preschoolers make a distinction between their beliefs and reality and that during the elementary school years, students become more aware of the role of their own cognitive processes (e.g., active attention, rehearsal, and organization) in facilitating their ability to perceive and remember events accurately (Wellman, 1990). Furthermore, a number of innovative curricular approaches have been devised for elementary school students—approaches that challenge students to think deeply about subject matter and to engage in authentic scientific inquiry as a community of learners (Brown & Campione, 1994; Hennessey, in press; Herrenkohl & Guerra, 1998; Lehrer, Carpenter, Schauble, & Putz, 2000; Metz, 2000; White, 1993). Students appear to thrive in these environments and can engage in impressive discourse about their ideas and the evidence for their ideas, at least with the scaffolding questions provided by their classroom teacher.

Thus, we believe the time is ripe to study what effects these more constructivist classroom environments have on elementary school students' epistemological understandings in science. It is important to know to what extent elementary students can begin to restructure their initial epistemological commitments and make progress in understanding theory, evidence, and the conjectural nature of scientific knowledge. It is also important to understand what kind of intermediate understandings may emerge in these environments. This work provides an initial study of the effects of one unusual elementary school curriculum on students' understanding of how knowledge is created and justified in science.

EPISTEMOLOGIES OF SCIENCE AS DOMAIN-SPECIFIC EPISTEMOLOGICAL THEORIES

Our framework for thinking about epistemological development is based on the assumption that even young children's concepts are organized in intuitive theories (Carey, 1985b; Gopnik & Meltzoff, 1997; Smith, Maclin, Grosslight, & Davis, 1997; Wellman, 1990) and that concepts in these theories undergo conceptual change (Carey, 1985b; Strike & Posner, 1985). This view assumes that young children can make use of abstract theoretical terms in generating explanations of everyday phenomena and that some of their concepts are organized in fairly coherent explanatory theories.

A leading example of such an early theory is preschoolers' theory of mind (Baron-Cohen, 1995; Gopnik & Meltzoff, 1997; Perner, 1991; Wellman, 1990), which Wellman argued is theory-like in three respects: (a) its concepts (e.g., belief, desire, mind, dream, imagine) are coherent and mutually interdefined, (b) these concepts support making fundamental ontological distinctions (i.e., between the mental and the physical), and (c) these concepts are part of an explanatory framework (i.e., they are used to explain the behavior of others and the origins of mental states). Note, too, that the concepts of belief, desire, and mind are not simple observables. Rather they are hypothetical constructs that help to explain many everyday observations. There is, of course, no claim that children of this age have a general concept of a theory or are aware of having a theory of mind. Nonetheless, they appear to make use of theory-like structures in their everyday thinking and reasoning.

Applying these ideas to epistemological development, we make the assumptions that children first develop a specific theory of knowing within the framework of their more general theory of mind between the ages of 4 and 6, and that they then go on to gradually complicate and restructure that theory based on their experience with generating, evaluating, and resolving competing knowledge claims in different domains. Because there are a number of domains in which elementary school students encounter competing knowledge claims (e.g., everyday life; school learning contexts for different subject areas such as math, science, and history), we assume that they can develop different epistemological stances in different domains. At the same time, the epistemological ideas students develop in one domain can inform, or provide resources for, restructuring epistemological ideas in other domains.

This view of epistemological development is similar to the views expressed by King and Kitchener (1994) and others in the general epistemological literature in that it assumes: (a) that epistemological views are coherent structures in which ideas about the nature of knowledge are related to ideas about how knowledge is justified and (b) that these views are qualitatively restructured during development. At the same time, it differs from these other views of epistemological development in three important respects. First, it does not assume there is some biologically based, general developmental constraint on epistemological development of the type proposed by either Inhelder and Piaget (1958) or neo-Piagetians such as Fischer (1980). Second, it situates elementary schoolchildren's initial epistemology as a subtheory within their emerging theory of mind, a theory that is thought to undergo significant development and elaboration during the elementary school years (Montgomery, 1992; Wellman, 1990). Finally, it assumes that children may develop different epistemological stances in different domains and that some aspects of epistemological standards are domain specific. Thus, it is important to develop tools for assessing epistemological understandings in specific domains to supplement tools that assess more general everyday epistemological commitments, such as the Reflective Judgment Interview (King & Kitchener, 1994). In this way one can study the development of epistemological commitments in different domains and examine how they may interact.

What specific concepts might inform students' epistemological commitments in the domain of science? Whereas the literature on everyday epistemological development uses general terms for describing the structure of knowledge (e.g., facts vs. views or viewpoints) and knowledge justification (e.g., authority-based vs. argument and evidence-based justification), more specific ideas are needed when describing epistemological commitments in science. Scientists do not just have views or viewpoints; they have specific theories that provide them with explanatory mechanisms and guide their generation of specific hypotheses. Furthermore, a distinctive and central aspect of their approach to knowledge justification is a process of designing experiments to test competing hypotheses.

Carey and Smith (1993) distinguished three qualitatively different epistemologies of science, each of which involves a set of different concepts for describing both the structure of scientific knowledge and the processes of knowledge acquisition in science.[2] Movement from one epistemology to another thus involves making fundamental conceptual changes.

At Level 1, scientific knowledge is assumed to consist of a collection of true beliefs about concrete procedures (e.g., how to do something correctly) or basic facts (e.g., what happens). Hence, students at this level make no clear distinction between scientists' ideas and activities or between their ideas and experimental results. They view scientific knowledge as accumulating in piecemeal fashion through simple telling or firsthand observation. They also view it as certain and true. Because they view scientific knowledge as about what to do and what happens, they view experiments as providing certain information about what happens or whether one's procedure works.

At Level 2, scientific knowledge is assumed to consist of a collection of tested ideas. The two new notions that emerge at this level are notions of explanation and hypothesis testing. Students at this level view scientists as concerned with under-

[2]In their recent book, Driver et al. (1996) also identified and characterized a progression of three qualitatively different epistemological views that they called (a) phenomenon-based reasoning, (b) relation-based reasoning, and (c) model-based reasoning. These views are similar in many respects to the three levels described by Carey and Smith (1993). We note, however, that the theoretical characterizations of Levels 1, 2, and 3 have been in a process of development, especially in terms of clarifying the distinction between Levels 2 and 3. Some aspects of Driver et al.'s middle view (relation-based reasoning, a simple inductivist epistemology in which there is no appreciation of scientists' concern with deep explanation or of the difficulties of acquiring knowledge) appears closer to what we would now describe as Level 1.5 than to Level 2. At Level 2, students are beginning to be aware that scientists develop hypotheses about underlying mechanisms and actively test these hypotheses via experiments.

standing how things work or why things happen. They also view scientists as doing experiments to test their ideas to see if they are right and as abandoning or revising their ideas when they find out they are wrong. Both notions of explanation and hypothesis testing require that students make a differentiation among scientists' ideas, activities, and experimental results (e.g., the purpose of an experiment is to test a scientist's idea; the purpose of an explanation is to account for an experimental result). Although students appreciate that prior knowledge influences the hypothesis-testing process, they still think absolute knowledge is obtainable with enough diligence and effort. They make no distinction between scientists' overarching theories and specific hypotheses.

At Level 3, scientific knowledge is believed to consist of well-tested theories about the world, which are useful in explaining events and predicting the outcomes of new events. A theory is understood as a coherent, explanatory framework that consists of a network of hypothetical theoretical entities that are used to explain patterns of data. Students at this level make an explicit distinction between the scientist's guiding theories and more specific hypotheses. They view theories as guiding all aspects of inquiry: the generation of hypotheses, the selection of methods, and the interpretation of data. They understand experimental results not only as providing evidence for and against hypotheses, but also as providing support (more indirectly) for and against theories. They also understand that theories, although revisable in principle, are resistant to change and slow to evolve. Ultimately, they judge canons of justification as framework relative and theories as more or less useful rather than strictly right or wrong. Thus, although students at Level 3 view scientific theories as providing rigorous standards for knowing and understanding, they also understand that knowledge of reality is fundamentally elusive and uncertain.

Carey and Smith (1993) noted that Level 1 ideas are consistent with what they dubbed a *knowledge unproblematic epistemology* (i.e., an epistemology in which knowledge is regarded as true and certain), whereas Level 3 ideas clearly reflect a *knowledge problematic epistemology* (i.e., an epistemology in which one understands the tentative, framework-relative nature of knowledge). These two contrasting epistemologies are very similar to the starting (absolutist) and ending (constructivist) epistemologies described in the general epistemological literature, although a Level 3 epistemology includes specific reference to the conjectural nature of explanatory theories and the role of indirect argument, evidence, and cycles of hypothesis testing in their evaluation.

Level 2 heralds the first emergence of some important domain-specific ideas (science as concerned with explanatory mechanism, experiments as a means for hypothesis testing) and represents one set of ideas that are transitional between these two epistemologies. On the one hand, a concern with explanation and testing does not immediately undermine one's belief in the true and certain nature of knowledge (e.g., one can believe that explanations are simple inductions from data, and that experiments can definitely prove one's hypothesis to be true). On the other, an acknowledgment that scientists are concerned with explanation and testing can sow seeds for appreciating the constructed, tentative nature of knowledge, especially as students begin to realize the conjectural nature of scientific explanations.

The Nature of Science Interview, developed by Carey and her colleagues (Carey, 1991; Carey et al., 1989), includes questions about the goals of science, the nature of scientific questions, the role of experimentation in science, the role of ideas in guiding experimentation, and the process by which scientists change their ideas. Each question is scored for whether it was answered in terms of Level 1, 2, or 3 ideas. Average level scores are then computed for each major section of the interview as well as across the whole interview.

To date, there have been two published studies with seventh-grade students (Carey et al., 1989; Honda, 1994) using this instrument. The findings have been similar in both studies: Seventh-grade students had average level scores of 1.0 across the interview. In addition, one study used the Nature of Science Interview with 11th-grade students (Honda, 1994). Although the majority of 11th-grade students gave Level 2 responses at some point in the interview, the overall average level score was 1.39, still closer to Level 1 than Level 2.

In these studies, students also were interviewed following a brief curriculum unit designed to develop Level 3 understandings. Students typically gained about one-half level. Thus, for 7th graders the average posttest level score was 1.5, and for 11th graders it was 1.89.

These studies did not directly test the notion that students with consistent Level 1 or Level 2 scores had multiple, different, mutually supportive Level 1 or Level 2 concepts. Further, the range of observed responses was very limited (mostly Level 1). In this study, we interviewed students who, because of their distinctive classroom experiences, may have more to say on these topics. We propose to elaborate on the coding system (to code for specific type of Level 1 or 2 concepts as well as for the general level of those concepts) to test the hypothesis that students at each level have distinct networks of mutually supportive ideas.

DEVELOPING A KNOWLEDGE PROBLEMATIC EPISTEMOLOGY: OPPORTUNITIES FOR CONCEPTUAL CHANGE IN THE ELEMENTARY SCHOOL CLASSROOM

Metz (1995) argued that the assumption that elementary schoolchildren are "concrete" thinkers has had a profound and limiting effect on the design of elementary school science curricula. Often, in the name of being developmentally sensitive, science curricula have focused on giving students practice with more concrete observation skills and ordering and classifying activities, rather than encouraging students to devise and test their emerging theories.

In recent years, a number of researchers, whose conception of the capabilities of elementary school students differs from the conception that informed the curricula just mentioned, have developed innovative elementary school curricular units that involve students in firsthand inquiry about important topics in science (Brown & Campione, 1994; Lehrer et al., 2000; Metz, 2000; White, 1993). Although these units differ from each other in numerous respects, all were developed with a shared commitment to a constructivist teaching pedagogy: a pedagogy that puts the primary focus on helping students to understand, test, and revise their ideas; a pedagogy that stresses the function of the social community in the negotiation of meanings and the growth of knowledge; and a pedagogy that gives students increasing responsibility for directing important aspects of their own inquiry. These researchers have documented the extent to which elementary schoolchildren are able to understand important explanatory ideas in scientific domains and are able to engage in meaningful inquiry and discourse about their inquiry in the scaffolded context of their classrooms. Yet no one has assessed via individual interviews whether these curricula have brought about fundamental changes in students' underlying epistemological views.

In this context, Hennessey's curricular approach stands out as an extensive and sustained attempt to teach elementary science from a coherent, constructivist perspective. As the sole science teacher for students in Grades 1 to 6, Hennessey had the opportunity to develop a curricular approach (over the past 20 years) that centers around helping students develop their own ideas about the world and how it works. Her approach works in three parallel areas—teaching for conceptual change, promoting student metacognitive understandings, and engaging students with deep domain-specific issues in science—and has been described by Hennessey (in press) herself and by others who have directly observed her classroom (Beeth, 1998; Beeth & Hewson, 1999a, 1999b).

Many interrelated pedagogical practices are part of her system of instructional design (Hennessey, in press). In describing her practice, Hennessey characterized the roles of the teacher, the student, the activities and tasks, and assessment. For example, as a teacher, she uses many of the conceptual change teaching strategies reported to be effective in the conceptual change literature, including (a) making students aware of their initial ideas and finding ways for students to make their ideas explicit, (b) encouraging students to clarify their ideas and to engage in metacognitive discourse about ideas, (c) employing "bridging analogies" and "anchors" to help students consider and manipulate new ideas (Clement, 1993; Clement, Brown, & Zietsman, 1989), (d) encouraging students to apply new understandings in different contexts, and (e) providing time for students to discuss the nature of learning and the nature of science. The students' role is to be actively involved in personal meaning making. This role requires that students take responsibility for representing their ideas, working to develop their ideas, monitoring the intelligibility of their ideas, considering the reasoning underlying specific beliefs, deciding on ways to test specific beliefs, assessing the consistency among their ideas, and examining how well these ideas extend to new situations. Hennessey selects activities and tasks based on their potential for engaging students in considering their own views, for helping students examine the reasoning that supports their views, and, at times, for promoting conceptual conflict or students' dissatisfaction with their views. Finally, assessment in her classroom is ongoing and multifaceted. Students are given many different options for communicating what they have understood. Such demonstrations involve students in presenting their findings in oral, written, visual, or graphic form; in raising new questions about a topic or idea; in applying their understanding to another context; or in reflecting on the growth of their understanding.

In a series of papers (Hennessey, 1994b, in press; Hennessey & Beeth, 1993), Hennessey also began to describe the impact of her approach on her students' thinking and understanding as revealed in the natural context of her classroom. Her main research focus has been on analyzing student writing and discourse to document the diverse ways her students have engaged in metacognitive reflection. These metacognitive acts range from identifying one's own beliefs and the beliefs of others to monitoring the intelligibility, plausibility, and fruitfulness of one's own beliefs and the beliefs of others. Data indicate that, by fourth, fifth, and sixth grade, her students have not only engaged in conceptual change themselves, but also have been able to learn metalanguage for describing the status of their changing conceptions.

Excerpts from class discussion and writing in which students responded to probes of their ideas about the learning process in science also revealed changes in these students' underlying epistemological ideas (Hennessey, 1995). First-grade students entered Hennessey's classroom with a simplistic epistemology (of the kind described in the literature, which focused on learning as the amassing of facts). By the end of the first grade, however, her students already were beginning to turn their attention to the central role of thinking in the learning process. Excerpts from students at other grade levels suggested that their views of thinking and learning processes in science were becoming more complex.

This study extends the effort to document the epistemological achievements of Hennessey's students by using the Nature of Science Interview (Carey et al., 1989). The study involves comparing the science epistemologies of Hennessey's sixth-grade students with the epistemologies of a demographically similar (White, middle class) group of sixth-grade students who also had an extensive elementary science curriculum taught by one teacher, but from a more traditional epistemological perspective. We included this comparison group to test the hypothesis that Hennessey's way of teaching leads to greater epistemological development than more traditional elementary science teaching. Carey's prior work (Carey et al., 1989) had been done with an urban, ethnically mixed, and socioeconomically diverse sample that was not strictly comparable in age or background to the students in Hennessey's classroom and the nature and extent of their elementary school science experiences were not documented. Hence it could not serve as an appropriate comparison group.

SPECIFIC HYPOTHESES

As discussed earlier, the Nature of Science Interview includes sections that probe student conceptions of the goals of science, the nature of scientific questions, the purpose of experiments, the role of ideas in scientists' work, and the nature of the processes by which scientific ideas change. We expected that students in Hennessey's classroom—who experienced a coherent constructivist pedagogy—would understand each of these issues differently than would students taught from a more traditional pedagogical perspective, reflecting pedagogical differences in the two classrooms. For example, in a constructivist pedagogy the main goals are to help students to pursue personal understanding and meaning making. In contrast, the main goals in a more traditional pedagogy are to help students to learn and memorize a large body of knowledge and find correct answers. In a constructivist pedagogy, the emphasis is on asking students deep domain-specific questions that help them to develop an understanding of the core explanatory ideas in a given domain. In contrast, greater emphasis is placed on asking students lower level factual and procedural questions in a more traditional pedagogy, and often the distinction between factual knowledge and underlying explanatory principles is blurred. Furthermore, in a constructivist pedagogy, students are given responsibility for deciding what experiments to do to test the ideas they hold. In contrast, students in more traditional classrooms typically are asked to carry out prescribed experimental procedures. These students, thus, may have little appreciation of the underlying purpose of an experiment and may not experience it as a test of ideas. Finally, students in constructivist classrooms experience the sustained development of their ideas by pursuing dialogue with peers about their ideas and engaging in multiple cycles of hypothesis testing and idea revision within a community of learners. In contrast, students in more traditional classrooms either simply read texts that provide them with a "rhetoric of conclusions" (Schwab, 1962) or carry out brief one-shot experiments after which they are presented with a correct conclusion or an answer to add to the other facts they have been asked to stockpile.

With these considerations in mind, we expected that the students who had participated in the more traditional curriculum would have developed coherent Level 1 views similar to those documented by Carey and her colleagues in previous studies. In contrast, we expected that students in Hennessey's classroom would have developed a different, richer network of concepts for thinking about knowledge construction in science—one centered on the more constructivist epistemology that characterized their curriculum. We expected that these students would see science as more centrally concerned with generating explanations and testing hypotheses (notions that first emerge at Level 2). We also predicted that students with a Level 2 epistemology would appeal to a rich set of new Level 2 ideas in their responses across the interview.

We were not sure how deeply Hennessey's students would conceptualize the notions of explanation and test. Would they go beyond the conceptions of explanation and test that are characteristic of Level 2 and begin to exhibit deeper Level 3 notions of framework theories, explanation in terms of unseen theoretical entities, and theory evaluation via complex cycles of hypothesis testing and revision to account for a wide pattern of data? Although we did not expect that these students would have a fully developed Level 3 epistemology (because that would call for more detailed scientific knowledge than these students had), we thought that they might have a more "enriched" Level 2 epistemology from what had previously been described, and that some students might be making progress in developing these difficult understandings. If elementary school students show an ability to develop a sophisticated, constructivist epistemology of science, then we would argue that elementary school can and should be a crucial time for developing and restructuring students' initial epistemological views.

METHODS

Design

This study involved two sixth-grade classes taught by different teachers in different schools. Each class had worked with the same science teacher throughout elementary school:

- Students in one classroom (henceforth called the constructivist classroom) were taught by Hennessey. She worked with the same students over a 6-year period, from first through sixth grade, with the students meeting three times a week for science class.
- Students in the comparison classroom had been with the same teacher for 5 years. They had science class five times a week in Grade 6, four times a week in Grade 5, three times a week in Grade 4, and once a week in Grades 1 through 3.

Our knowledge of the curriculum in the constructivist classroom is gleaned from conversations with Hennessey, research papers analyzing her classroom environment, and our own classroom observations. Our knowledge of the curriculum in the comparison classroom is based on a 2-hr interview with the science teacher, a separate 2-hr meeting with the principal, and inspection of the science facilities and classroom materials.

Hennessey's science curriculum centers on engaging students' own ideas, with the teacher taking the role of facilitator of both small-group work and whole-class discussion. Students typically work in groups of four to investigate phenomena and to develop their own personal models for explaining these phenomena. As students engage in experimentation and dialogue with each other and with Hennessey, she guides them to reflect on the intelligibility, plausibility, and fruitfulness of their ideas. The class functions as a community of learners where students engage in a process of trying to make sense of their own and each other's ideas. Even in first grade, students invent ways to make their ideas understandable to others. Reading in this classroom occurs when students seek information from books to enhance their explorations and to help them make sense of their ideas. The evolution of the curriculum is open ended in nature, with the teacher suggesting time for investigation or sharing in direct response to issues, questions, or obstacles that arise in small-group work.

Both the students' and teacher's roles change as the students become more metacognitively sophisticated from Grades 1 to 6 (see Table 1). For example, in first grade, the goal is for students to state explicitly their own views about the tasks under consideration and to begin to generate reasons that support these views. That is, the aim for students is to begin to differentiate what they think from why they think it. The teacher's role in this endeavor is to provide many experiences and opportunities in which students can begin to articulate the reasoning used to support their views and to help students find a variety of ways to represent their thinking externally. By Grades 4 through 6, students not only differentiate what they think from why they think it, but also become explicitly aware of a variety of criteria by which their ideas can be evaluated; for example, consistency with other ideas, fit with a pattern of evidence, and fruitfulness in explaining new phenomena. They also actively monitor the status of their own thinking vis-à-vis its intelligibility, plausibility, and fruitfulness. The teacher uses the information students provide on the status of their thinking to guide her in challenging their thinking, in introducing them to new ideas, and in engaging them in the process of conceptual change.

Hennessey selects curricular content that will engage students with fundamental scientific questions and optimize the reaching of metaconceptual goals. For example, first-grade students participate in activities and work with topics that successfully engage them in explanatory model building (e.g., the day–night cycle). Curricular units start with students exploring or producing phenomena during a process of "messing about." Children are invited to suggest what they make of a given phenomenon and to decide (in their small groups) how to proceed to see which of their ideas about the phenomenon are most plausible. At several points in this cyclic process they report to their classmates, who, in turn, carefully review their arguments. As the curriculum progresses, students assume more control of the topics studied, with the emphasis on explanatory understanding and theory building.

TABLE 1
Changes in the Students' and Teacher's Pedagogical Roles Across Grades 1 Through 6 in the Constructivist Classroom: A Progression of Increasingly Sophisticated Metaconceptual Activities

Grade	*Students' Roles*	*Teacher's Roles*
1	•Explicitly state their own views about the topic under consideration. •Begin to consider the reasoning used to support their views. •Begin to differentiate what they think from why they think it.	•Finds a variety of ways in which students can externally represent their thinking about the topic. •Provides many experiences for students to begin to articulate the reasoning used to support ideas and beliefs.
2	•Begin to address the necessity of understanding other (usually peer) positions before they can discuss or comment on those positions. •Toward end of the year, begin to recognize inconsistency in the thoughts of others, but not necessarily in their own thinking.	•Continues to provide an educational environment in which students can safely express their thoughts without reproaches from others. •Introduces concept of consistency of thinking. •Models consistent and inconsistent thinking (students can readily point out when the teacher is being inconsistent!).
3	•Explore the idea that thoughts have consequences, and that what one thinks may influence what one chooses to see. •Begin to differentiate understanding what a peer is saying from believing what a peer is saying. •Begin to comment on how their current ideas have changed from past ideas and to consider that current ideas may also need to be revised over time.	•Fosters metacognitive discourse among learners to illuminate students' internal representations. •Provides examples from their personal work (which is saved from year to year) of student ideas.
4–6	•Begin to consider the implications and limitations of their personal thinking. •Begin to look for ways of revising their personal thinking. •Begin to evaluate their own and others' thinking in terms of intelligibility, plausibility, and fruitfulness of ideas. •Continue to articulate criteria for acceptance of ideas (i.e., consistency and generalizability). •Continue to employ physical representations of their thinking. •Begin to employ analogies and metaphors, discuss their explicit use, and differentiate physical models from conceptual models. •Articulate and defend ideas about what learning should be like.	•Provides historical examples of "very important people" changing their views and explanations over time. •Begins to use students' external representations of their thinking as a way of evaluating their ideas and beliefs (in terms of intelligibility, plausibility, and fruitfulness) to: (a) create, when necessary, dissatisfaction in the minds of the learner to facilitate conceptual exchange; or (b) look for ways of promoting conceptual capture within the mind of the learner.

The comparison classroom's science curriculum, which also spans Grade 1 through Grade 6, centers on a more traditional, knowledge unproblematic approach. In the first, second, and third grades the teacher presents a topic and the students focus on learning facts and on creating art about that topic (i.e., drawing, cutting, and pasting). If the topic lends itself to everyday experience, they may explore outdoors or at home with hands-on activities. In the fourth, fifth, and sixth grades, topics in science are presented by the teacher in a lecture format and the students are assigned corresponding readings in a standard text.

The fifth and sixth graders also participate in an annual school-sponsored science fair, for which they choose a topic, make hypotheses, design and execute experiments, and present their results to the public at an evening event. The teacher regards this as one of the most important, culminating events in the science curriculum, which encourages students to think for themselves, generate a hypothesis and a method, carry out an experiment, and draw conclusions. Examples of some of the kinds of science fair questions her students have investigated include: What kind of wood burns the longest? Which brand of paint is most adhesive after weather conditions? These questions focus on whether there is a relation between two observable variables that are part of students' everyday experience and are the type students easily can investigate in a simple experiment that provides straightforward answers. They involve a personally meaningful topic and provide a reason for learning about some aspects of scientific method. Note, however, that they are not deep explanatory questions about how things work, which appeal to unseen theoretical entities; they do not involve students with sustained investigations of phenomena, where they pursue answering a given question in multiple ways.

The comparison classroom teacher's main aims are for students to like science and think that it connects to their everyday lives. She offers a variety of approaches to learning, including watching videotaped television programs on science, discussing science topics in the news, and engaging in group discussions, hands-on experiments, drawing, and outdoor activities. She is concerned that students be successful in their classroom endeavors (even poor readers, whom she perceives to be at a disadvantage). The teacher selects curricular content based on what is most interesting to the students, what has been successful in past years at particular ages, and what can be accomplished within the workspace and time available. Overall, her own epistemology of learning in science seems to be based predominantly on notions of problem solving and critical thinking rather than notions of theory building in a community of learners, sustained metacognitive development, and conceptual change. Furthermore, many of her self-reported pedagogical practices involve students in factual learning (i.e., reading from texts) or concrete problem solving (i.e., trying to put together an electric circuit that will work) rather than in the developing and testing of deep explanations for the workings of real-world phenomena.

Population

The two schools served similar student populations (White, middle to upper middle class). Both were private Catholic coeducational schools that, according to school staff, were selected by parents because of their good reputations. Class sizes were comparable. Hennessey had one sixth-grade class of 22 students. The comparison school had two sixth-grade classes that contained a total of 36 students; each class had the same science curriculum and science teacher. All sixth-grade students in both schools were sent permission letters for participating in our interviews. Eighteen students from the constructivist classroom and 27 students from the comparison classrooms accepted the invitation and were interviewed.

Assessment

In a private, one-on-one setting, all participants received the Nature of Science Interview developed in prior work by Carey (Carey, 1991; Carey et al., 1989). Because this interview instrument was developed prior to our work with these classes, it was not tailored in any way to mirror the language of either classroom. The interview took about 20 to 30 min during which students responded to direct questions about the scientific enterprise. The interview was tape-recorded and transcribed for analysis. The script of this interview is included in Appendix A.

Analysis

We undertook two main analyses of the Nature of Science Interview: (a) an analysis of the type and the level of ideas expressed in four separate clusters of interview questions (question cluster analysis), and (b) an analysis of the consistency and the coherency of ideas expressed by individual students across the four question clusters (consistency and coherency analysis). In addition, we supplemented these analyses with several holistic analyses of students' entire interview.

The question cluster analysis involved our reading the entire corpus of students' verbatim transcripts separately for each of the four question clusters (i.e., goals of science, types of questions, nature and purpose of experiments, and nature of change processes). We were blind to students' identities and classrooms when developing coding categories to capture the type and level of ideas that were present in a given cluster. In developing these coding categories, we were responsive both to important issues in Carey's levels system and to what emerged more inductively from the data for those questions.

Key ideas that reflect a Level 1 knowledge unproblematic epistemology focus on science as involving (a) simple activities and procedures (e.g., trying things out, doing experiments) or (b) acquiring factual knowledge (e.g., learning new things, making discoveries). In such an epistemology, knowledge is acquired in a piecemeal and unproblematic fashion by making observations and doing experiments, and students do not differentiate clearly between scientists' ideas, experimental methods, and experimental results.

In contrast, four key ideas that reflect a Level 2 epistemology focus on science as concerned with (a) explaining how things work or why things happen, (b) testing hypotheses or prior ideas, (c) developing ideas, and (d) working to understand these ideas. Simple expressions of these ideas (e.g., one of these ideas expressed alone in a question cluster) were scored as Level 2. More complex and sophisticated expression of these ideas (e.g., two of these ideas expressed within the same question cluster) were scored as Level 2.5 (see Appendix B). In general, a Level 2.5 epistemology involves the same basic conceptual distinctions as Level 2, but the concepts are more elaborated and explicitly interrelated.

A Level 3 score is based on understanding that theories are coherent explanatory frameworks that guide the scientists' construction of specific hypotheses, design of experiments, and interpretation of experimental evidence. Theories are seen as inherently conjectural in nature and are developed through sustained cycles of hypothesis testing, revision, and indirect argument from patterns of data. An important conceptual distinction is made between specific hypotheses and the general framework from which they are derived. This differentiation between theories and hypotheses allows the formulation of explicit relations between these two concepts and the understanding that experiments provide evidence not only for or against specific hypotheses, but also for or against the larger framework theory. (Note that no response in this study warranted a Level 3 score.)

Appendix B provides a detailed description of the actual coding categories used for each question cluster. It specifies the type and level of a given coding category (see italicized terms), provides examples of typical responses for that category, and gives reasons the category is situated at a given level. It also specifies the ways Level 2.5 answers fall short of Level 3 responding. Note that this coding system contains a transitional Level 1.5 between Level 1 and 2. Included within this level are some responses that are ambiguous and could be coded as either Level 1 or Level 2 (e.g., mention of how things work, unsupported by a particular example). Also included are responses that seem more sophisticated than classic Level 1 responses but fail to meet the criteria for Level 2 responses (e.g., questions about how two observable variables relate are more sophisticated than simple procedural or factual questions, but not as sophisticated as explanatory or theoretical questions).

Each question cluster was scored independently by two coders. Because a given student could express several different ideas within a cluster, coders checked off all the coding categories that applied. Intercoder agreement on the positive occurrence of each type of level code listed in Appendix B was computed (i.e., number of times both coders checked the particular code divided by number of times at least one coder checked that code). Agreement averaged 80% and ranged from 67% to 96% for different codes. This calculation is conservative because it ignores agreements that a given type of code was absent for a given student. The overall agreement on the presence or absence of a specific type of code was even higher, averaging over 90%. Disagreements were resolved through discussion.

Students were then assigned a level score for each question cluster. In general, the level score was simply the highest level idea code assigned. Students were given a level score of 2.5 based on the presence of multiple Level 2 ideas within a cluster or the presence of a sophisticated Level 2.5 idea (see Appendix B for a description of the criteria for Level 2.5 for each cluster). To determine the reliability of level codes, we computed level scores based on each coder's pattern of codes for a given student. Agreement on level scores was 82% for Cluster 1, 87% for Cluster 2, 89% for Cluster 3, and 82% for Cluster 4.

RESULTS

The main analyses focus on (a) the type and level of ideas expressed by students in the four main question clusters, and (b) the pattern of ideas across the four question clusters. If students in the two classrooms have different epistemological stances, then we would expect that they would appeal to different ideas in the four question clusters and that they would show consistency and coherency in the ideas expressed. This expectation grows out of our assumption that a network of distinctive and mutually supporting ideas characterize a particular epistemological stance.

The main analyses are followed by several supplementary analyses that were conducted by reading each student's interview as a whole. The supplementary analyses were carried out as further tests of whether contrasting epistemological stances exist for the two classrooms.

Question Cluster Analysis

Cluster 1: Goals of science. Students mentioned several different ideas in response to questions about the goals of science: (a) doing things (Level 1), (b) gathering information (Level 1), (c) thinking about ideas or data (Level 1.5), (d) finding explanations (Level 1.5 for unelaborated or ambiguous mention of how it works, and Level 2 for discussion of why it happens and elaborated discussion of how it works), (e) testing ideas (Level 2 or 2.5), (f) understanding ideas (Level 2), and (g) developing ideas (Level 2 or 2.5). Figure 1 displays the percentage of students in each classroom who mentioned each kind of idea. Two-tailed *t* tests were

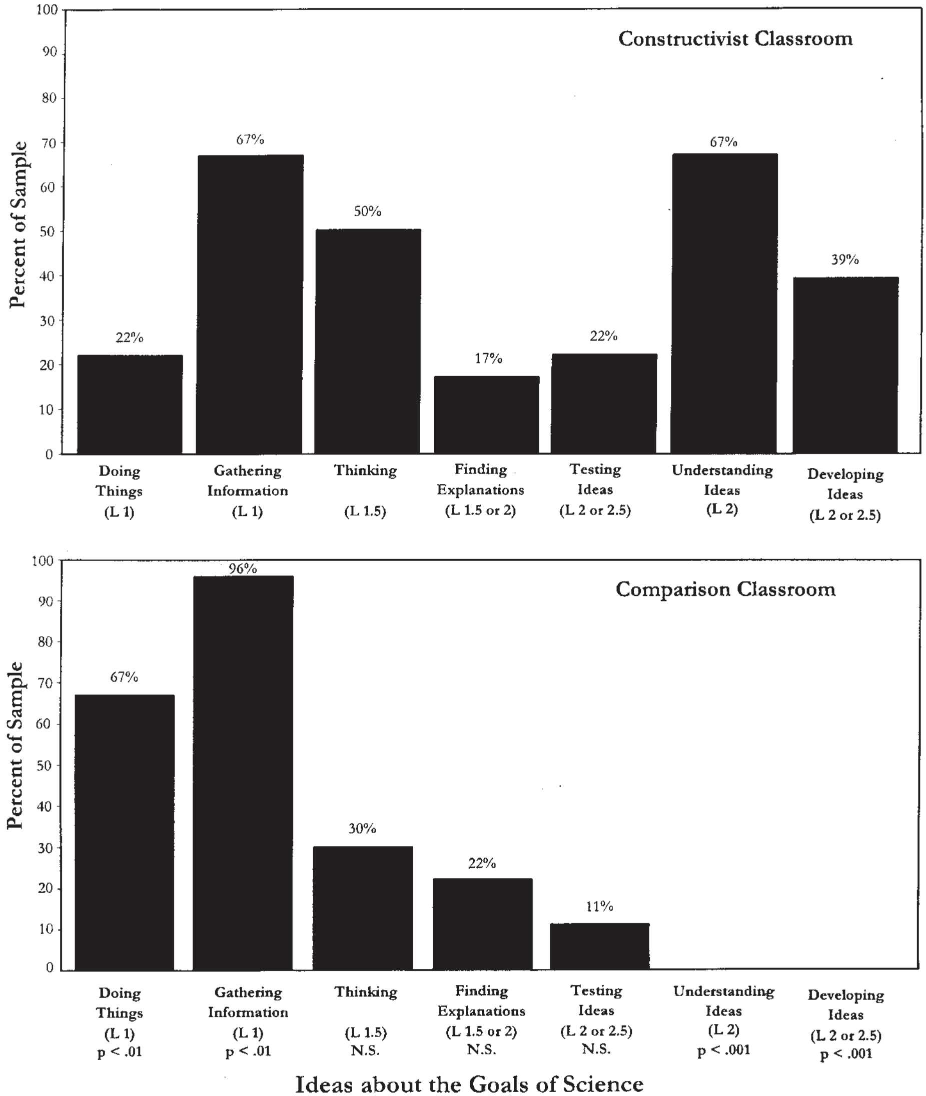

FIGURE 1 A comparison of the percentage of students expressing different ideas about the goals of science in the constructivist and comparison classrooms. L1, 1.5, 2, and 2.5 indicate Level 1, 1.5, 2, and 2.5 ideas, respectively. Significance level for each comparison is given at the bottom of the figure.

performed to indicate for which kinds of ideas the classroom differences were statistically significant.[3]

The ideas mentioned by the greatest percentage of students in the comparison classroom were the Level 1 notions of doing things and gathering information. At Level 2, fewer mentioned finding explanations or testing ideas and none mentioned understanding ideas or developing ideas.

There was a very different pattern for students in the constructivist classroom. The Level 2 ideas mentioned by the greatest percentage of students were understanding ideas and developing ideas (ideas that had not been mentioned at all by students in the comparison classroom). These processes are central, framing concepts in a constructivist epistemology and were central practices in this classroom. It is striking that they are often the first ideas mentioned by these students. Fewer students mentioned the Level 1 ideas of doing things (with no articulated purpose) and gathering information than in the comparison classroom, although many still mentioned the idea of gathering information. As in the comparison classroom, these opening questions drew little mention of finding explanations or testing ideas.

Each student was assigned a level score based on his or her highest level idea expressed in this question cluster. Students who either related two or more Level 2 ideas to the goal of science or discussed a sophisticated Level 2.5 response received a level score of 2.5. All students in the constructivist classroom who received a score of Level 2.5 saw the goal of science as both understanding their ideas and either developing or testing them (typically through some multistepped process). The 1 student who received a Level 2.5 in the comparison classroom saw the goal of science as figuring out how things worked and testing one's ideas.

The mean level score was 2.1 for students in the constructivist classroom compared to 1.4 for students in the comparison classroom, a significant difference (t = 5.36, $p < .0001$, two-tailed). Individual scores for students in the constructivist classroom ranged from 1.5 to 2.5, with a modal score of Level 2. Individual scores for students in the comparison classroom ranged from 1 to 2.5, with a modal score of Level 1.

Cluster 2: Type of questions. In this question cluster, students were asked to give an example of the type of questions scientists ask. We classified the ques-

[3]Note that, in Figures 1 through 4, we present t tests of classroom differences in the percentage of students that mention each kind of idea only as a means of enhancing our description of the findings. We realize that with repeated t tests a few differences would be expected to be significant by chance. We thus think it is important to emphasize here that our main hypotheses concerned differences in average level scores for each question cluster and that our test of each hypothesis involved one t test per cluster (presented at the end of each section).

tions they discussed into six main types: (a) procedural questions (Level 1), (b) journalistic questions (Level 1), (c) variable relation questions (Level 1.5), (d) explanation questions (Level 1.5 for unelaborated questions about how it works, Level 2 for questions about why something happens or for an elaborated discussion of how something works), (e) questions involving some theoretical entity (Level 1.5 when treated more factually, Level 2 when not), and (f) metacognitive questions (Level 1.5 when about Level 1 issues, Level 2 when about Level 2 issues). Figure 2 contrasts the percentage of students in the two classrooms who discussed each type of question.

In the comparison classroom, the majority mentioned journalistic questions (i.e., questions about observable events or concrete objects: who, what, where, and when questions). In addition, one third of the students mentioned procedural questions (i.e., how to do something) and 30% mentioned variable relation questions (e.g., If basketball teams play at home or away does it affect how many points they score? That is, is there a home-court advantage? Does listening to music affect how quickly you do your homework?). The latter questions were identified by students as the questions they had investigated for their science fair projects at school. Note that these classic science fair questions focus on investigating whether an association exists between two easily measured variables, rather than probing for deeper explanations of the association. All three question types are consistent with a knowledge unproblematic epistemology and were more frequent in the comparison classroom than in the constructivist classroom. Forty-one percent of the students in the comparison classroom also gave explanation questions, which ranged from ambiguous Level 1.5 questions about how something works to clearer Level 2 questions about why something happens.

In contrast, the majority of students in the constructivist classroom focused on explanation, metacognitive, and theoretical entity questions. Explanation questions are a classic Level 2 type of question in Carey's levels analysis (Carey et al., 1989). Metacognitive questions is a new category we developed in light of this data set. Students imagined scientists asking themselves both basic goal-oriented and activity-oriented Level 1.5 metacognitive questions, such as "What am I trying to accomplish?" and "Why am I doing this experiment?", as well as more sophisticated Level 2 metacognitive questions involving the clarity and reasons for one's ideas, such as "Why do I think that (i.e., what are my reasons)?" "How intelligible or plausible are my ideas?"—types of metacognitive questions Hennessey explicitly encouraged students to ask from fourth through sixth grade (see Table 1). Questions involving theoretical entities refer to unobservable, theoretical entities rather than to concrete objects and observable phenomena. Examples include questions about atoms, DNA, gravity and other forces, and germs. Again, concern with understanding the deep domain issues of the nature of matter, force and motion, and the workings of the human body were all a central part of Hennessey's curriculum.

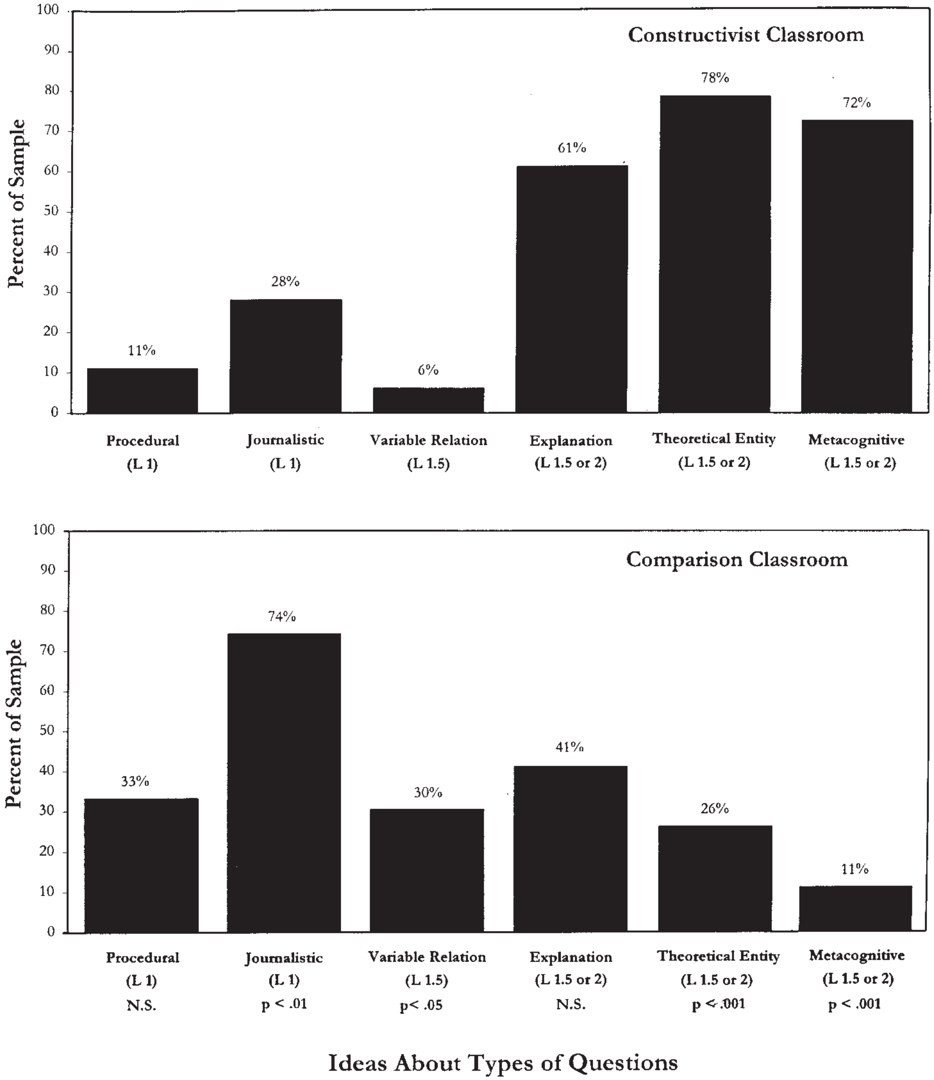

FIGURE 2 A comparison of the percentage of students expressing different types of questions in the constructivist and comparison classrooms. L1, 1.5, and 2 indicate Level 1, 1.5, and 2 ideas, respectively. Significance level for each comparison is indicated at the bottom of the figure.

Both metacognitive and theoretical entity questions were much more common in the constructivist than in the comparison classroom. Furthermore, students in the constructivist classroom generally asked more sophisticated explanation, metacognitive, and theoretical entity questions than students in the comparison classroom. In the comparison classroom, almost half of the explanation questions were of the unelaborated "how it works" type; all of the metacognitive questions focused exclusively on scientists' goals and activities; and more than half of the questions involving theoretical entities were stated as simple factual questions. In the constructivist classroom, all of the explanation questions were eventually elaborated with a specific example (e.g., Why do we get diseases? or, Why does the earth go around the sun?). Half of the metacognitive questions raised by students concerned the intelligibility, clarity, or reasons for their ideas; and all but one question involving a theoretical entity were treated as ambiguous or more complex to answer than a simple factual question.

Students were assigned a level score based on the highest level question they expressed in this cluster. Students were given a Level 2.5 if they combined two or more Level 2 codes. Most of the students in the constructivist classroom with Level 2.5 scores combined an explanatory question with some reference to a theoretical entity (e.g., How do you think this works? [Interviewer: Can you give an example?] …. Do you think electrons circle around the atom?; or, How does it move? How does it work? What motions are there? What kind of motion? What kind of forces?). This type of combination makes it clear that a student is considering a deeper level of explanation—one concerned with unseen entities, not just observable ones. (This concern with deeper explanation is generally implied at Level 2, but is not made as explicit.) Several students in the constructivist classroom combined a specific explanatory or theoretical question with a sophisticated question about the intelligibility of their ideas. The fact that they realized their ideas take work to understand provides us with indirect evidence that these ideas are not about simple observables.

The mean level score for students in the constructivist classroom was 2.3, compared to 1.5 for students in the comparison classroom, a significant difference (t = 5.62, $p < .0001$, two-tailed). Individual level scores for students in the constructivist classroom ranged from 1.5 to 2.5, with a modal score of Level 2.5. Individual level scores for students in the comparison classroom ranged from 1 to 2.5, with a modal score of Level 1.5.

Cluster 3: Nature and purpose of experiments. In this question cluster, students were asked first "What is an experiment?" and then "Why do scientists do experiments?" The five main kinds of ideas we coded were (in order of increasing sophistication): (a) try out or find cures (Level 1), (b) find answers (Level 1), (c) find explanations (Level 1.5 for an unelaborated mention of how it works, otherwise Level 2), (d) test ideas (Level 2), and (e) develop ideas (Level 2 when alone, Level 2.5 when in combination with test ideas). Figure 3 contrasts the percentage of students in the two classrooms mentioning each of these ideas.

Overwhelmingly, students in the comparison classroom saw experiments as a means of trying things out or finding cures and finding answers, both clear knowledge unproblematic ideas. Students in the constructivist classroom did mention these ideas, but less commonly. Furthermore, the notions of test ideas and develop ideas were very strong for students in the constructivist classroom, whereas little or no mention of these ideas occurred in the comparison classroom. These two notions are consistent with a more constructivist epistemology. We suspect that their strong showing in this classroom was supported by the students' long-term experience with designing ways of testing their ideas.

Students were assigned a level score based on the highest level idea they expressed in this question cluster. Students who related two or more Level 2 ideas to the purpose of experiments received a 2.5. Invariably, the two ideas combined by students in the constructivist classroom were the notions of testing and developing ideas. That is, these students not only mentioned that the purpose of experiments was to test a specific idea, but also went on to talk about how experiments were important in the larger enterprise of trying to develop one's ideas (i.e., make them better, clearer to oneself and others, etc.). In this way, the experimental testing of ideas was seen not as a limited process of finding out if an idea is right or wrong, but as part of a broader goal that is central to a constructivist epistemology: deepening ideas and one's understanding.

The mean level score for students in the constructivist classroom was 2.0 compared to 1.2 in the comparison classroom, a significant difference ($t = 4.59$, $p < .0001$, two-tailed). Individual scores for students in the constructivist classroom ranged from 1 to 2.5, with a modal score of Level 2.5. Individual scores for students in the comparison classroom ranged from 1 to 2, with a modal score of Level 1.

Cluster 4: Nature of change processes. The last cluster of questions probed for student understanding of the conditions that lead scientists to change their ideas and theories. Again, there were differences in the level of sophistication of student responses in the two classrooms. In scoring these questions, we identified six key ideas. The first idea reflected an extremely simple view of the change process: Scientists easily keep or abandon an idea based on whim, or based on a single observation or experiment (Level 1). Other students were beginning to realize the change might require more thought or effort (Level 1.5), although they did not yet articulate distinctive Level 2 ideas. Finally, other students mentioned distinctive Level 2 ideas, which revealed a more sophisticated view of the change process. These four ideas were that (a) change involves development of ideas (rather than simply keeping or abandoning them), (b) change involves complex evidence (rather than only one straightforward observation), (c) change involves finding a better explanation (rather

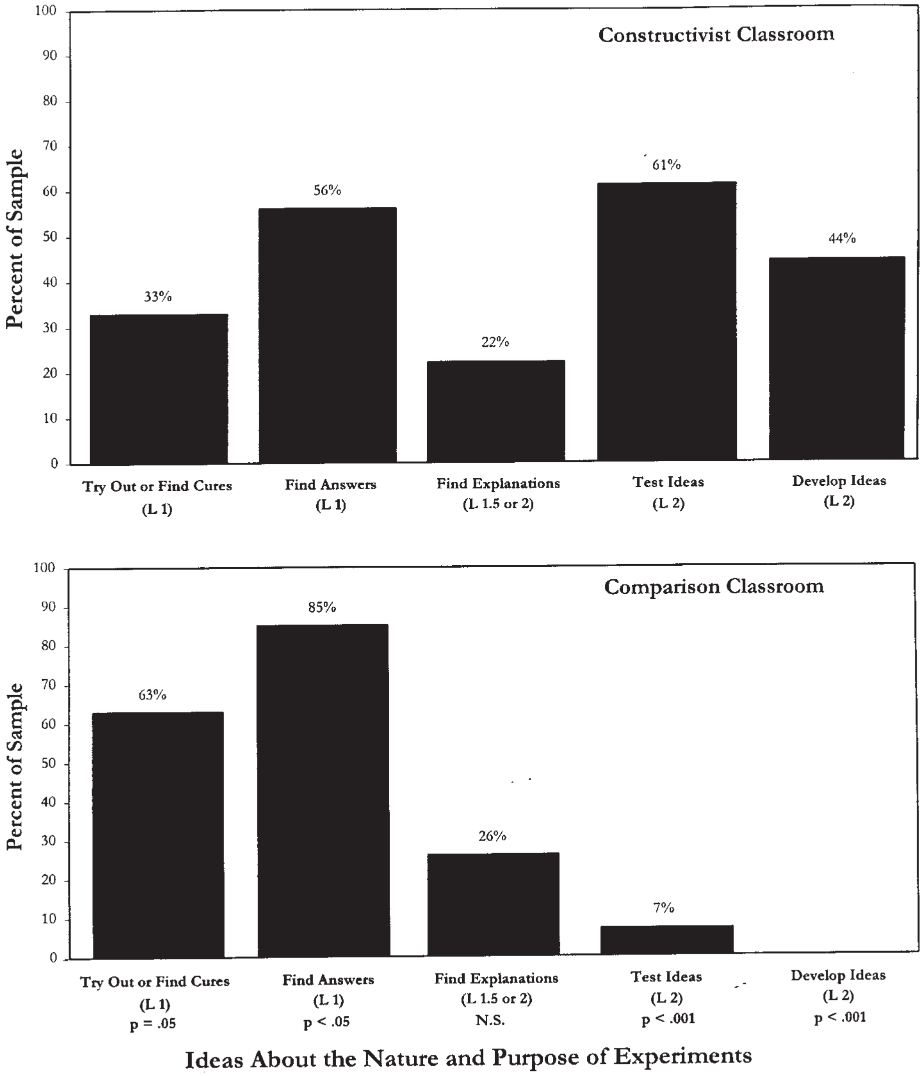

FIGURE 3 A comparison of the percentage of students expressing different ideas about the nature and purpose of experiments in the constructivist and comparison classrooms. L1, 1.5, and 2 indicate Level 1, 1.5, and 2 ideas, respectively. Significance level for each comparison is indicated at the bottom of the figure.

than finding new facts or answers), and (d) change is constrained by prior ideas (rather than simply reflecting new observations). Figure 4 contrasts the percentage of students in the two classrooms who mentioned each of these ideas.

The dominant view (expressed by more than half of the students) in the comparison classroom was that scientists keep or abandon an idea after a simple observation or a single experiment—a view consistent with a knowledge unproblematic epistemology. One third of the students mentioned that thought or effort was involved in the change process. Fewer mentioned any of the Level 2 ideas. Those that did mention Level 2 ideas typically talked about change as involving the development of ideas (15%) or indicated, by their example, that the change involved a better explanation (11%). None made a metastatement that change involved finding better explanations. Only one mentioned complex evidence and none discussed the fact that change is constrained by prior ideas. Two students (7%) mentioned two different Level 2 ideas.

In contrast, no student in the constructivist classroom expressed only the simple keep or abandon view of the change process. Indeed, mention of diverse Level 2 ideas was abundant. Sixty-one percent mentioned that change involves the development of ideas. In addition, some mentioned that change involves complex evidence (39%), that it involves finding a better explanation (33%), or that it is constrained by prior ideas (39%). Forty-four percent of students mentioned two or more Level 2 ideas.

Students were then assigned a level score based on their highest level idea expressed in this question cluster. Students who saw change as constrained by both the Level 2 factor of fit with evidence from multiple experiments (complex evidence) and the Level 2 factor of fit with prior ideas (metastatements by students that change is constrained by prior ideas or that change occurs when one finds a better explanation) were scored as Level 2.5. These students were moving beyond some simple notion of hypothesis testing (i.e., finding out if an idea is right or wrong) to realizing that ideas have to be evaluated based on both their fit with evidence and with other ideas.

The mean level score for students in the constructivist classroom was 2.1 compared to 1.3 in the comparison classroom, a significant difference ($t = 6.98$, $p < .0001$, two-tailed). Individual scores for students in the constructivist classroom ranged from 1.5 to 2.5, with a modal score of Level 2. Individual scores for students in the comparison classroom ranged from 1 to 2, with a modal score of Level 1.

Consistency and Coherency Analysis

Consistency of levels. We averaged students' scores for the four question clusters to obtain a measure of the consistency of their Level 1 versus Level 2 responding. Figure 5 shows the contrasting distributions for students in each classroom. Again there were striking differences in average level scores: The majority

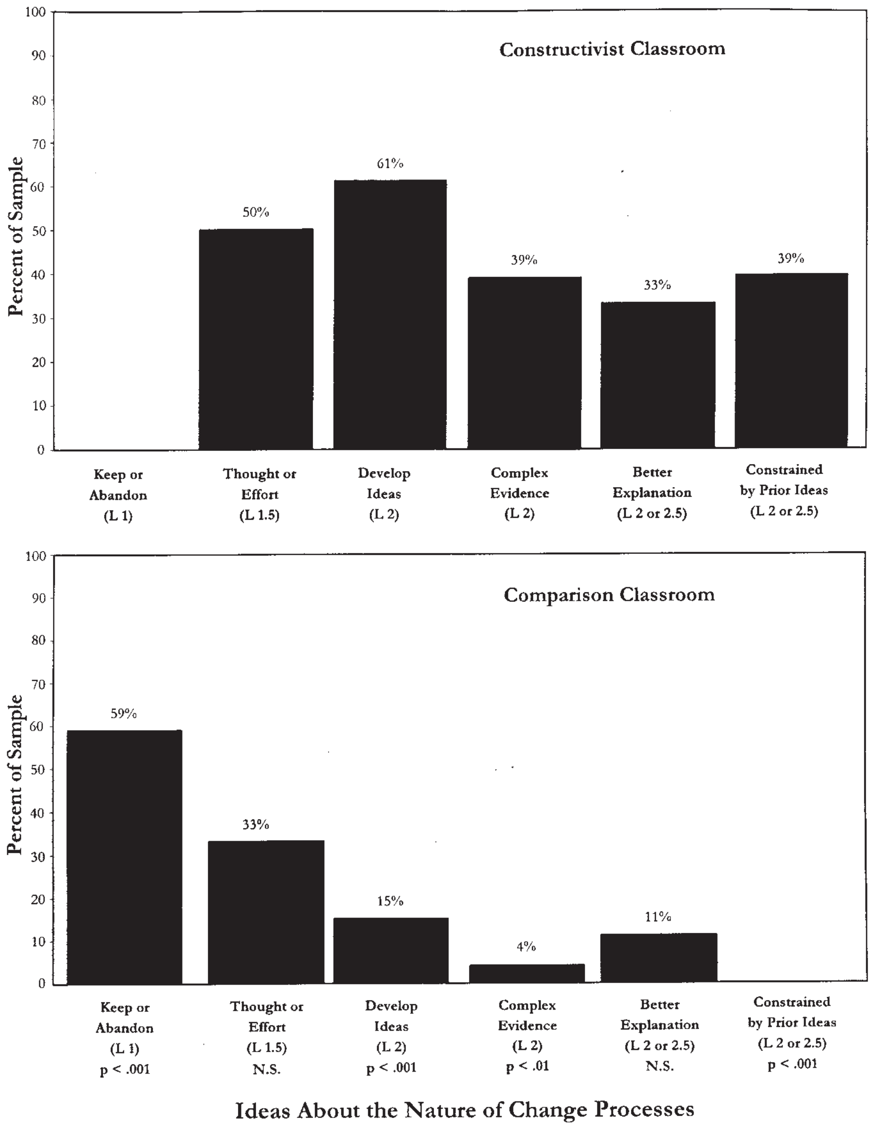

FIGURE 4 A comparison of the percentage of students expressing different ideas about the nature of change processes in the constructivist and comparison classrooms. L1, 1.5, 2, and 2.5 indicate Level 1, 1.5, 2, and 2.5 ideas, respectively. Significance level for each comparison is indicated at the bottom of the figure.

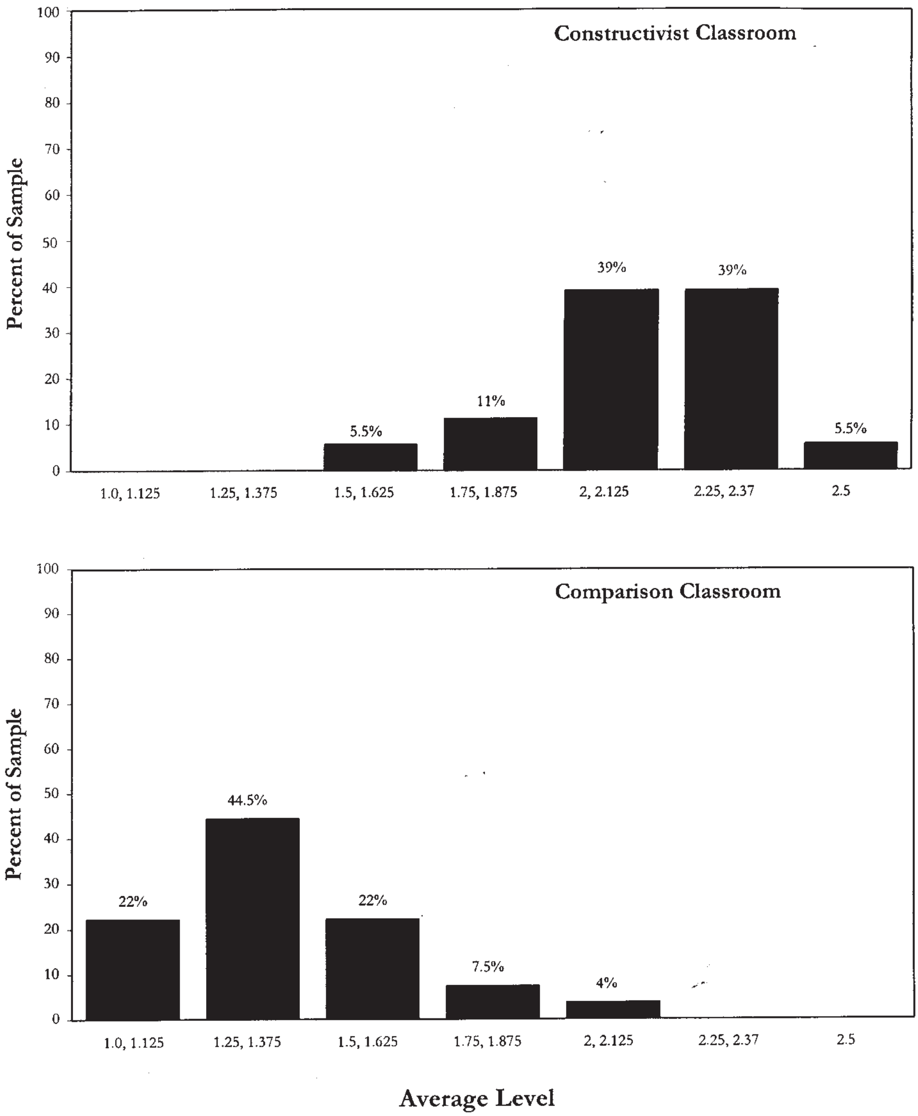

FIGURE 5 A comparison of the distribution of average level scores (level scores averaged across the four question clusters) in the constructivist and comparison classrooms. Given our scoring, 13 average values were possible, ranging from 1.0 to 2.5. For simplicity of presentation, adjacent values (e.g., 1 and 1.125, or 1.25 and 1.375) are grouped together.

of students in the constructivist classroom had an average score of Level 2 or above, whereas the majority of students in the comparison classroom had an average score between 1 and 1.37. There was very little overlap in the two distributions of scores.

Coherence of Level 2 ideas. It is possible that a student could show consistent Level 2 reasoning by appealing to the same Level 2 idea repeatedly across each question cluster. However, if consistent Level 2 reasoning involves a reorganization of student ideas around a new set of Level 2 concepts, there should be a coherent network of Level 2 concepts that mutually support each other and emerge as part of this restructuring. Indeed, new concepts take on meaning as constructivist ideas only in the company of other constructivist ideas.

In this analysis, we grouped the Level 2 responses from the preceding question clusters analysis according to whether they were concerned with understanding ideas, explanation, testing ideas, or developing ideas.[4] We then plotted the average number of different Level 2 ideas mentioned by students as a function of their average level score across the four clusters.

Figure 6 shows that there was a strong relation between students' average level score and the number of different Level 2 ideas they mentioned and that the relation was similar for students in the two classrooms. Students who mentioned only one Level 2 idea were generally still entrenched in Level 1 thinking. As students moved toward more consistent Level 2 responding, they also increased the number of different Level 2 ideas they mentioned. Students with average level scores of 1.75 or 1.87 generally expressed two different constructivist ideas. Those with average level scores of 2 or 2.125 generally expressed three different constructivist ideas. Those with average level scores of 2.25 and above generally expressed four different constructivist ideas. Thus, consistency in responding across a variety of questions reflects the emergence of a network of new ideas that mutually support each other.

Modal epistemology in the constructivist classroom. These four Level 2 ideas—understanding ideas, explanation, testing ideas, and developing ideas—describe processes that were central to the modal epistemology of students in the constructivist classroom. Fifty-five percent of the students in the

[4]These four distinct Level 2 ideas emerged in the specific coding categories of the different question clusters. Specific coding categories grouped as manifestations of understanding ideas were: (a) understanding ideas (Cluster 1), (b) metacognitive questions about Level 2 issues (Cluster 2), and (c) need for new ideas to make sense or need for fit and coherence among ideas (two subcategories of constrained by prior ideas, Cluster 4). Specific coding categories grouped as manifestations of explanation were: (a) Level 2 explanation ideas (Clusters 1, 2, 3, 4), and (b) Level 2 theoretical entity questions (Cluster 2). Specific coding categories grouped as manifestations of testing ideas were: (a) testing ideas (Clusters 1, 3), and (b) complex evidence (Cluster 4). Specific coding categories grouped as manifestations of developing ideas were the develop ideas codes for Clusters 1, 3, and 4.

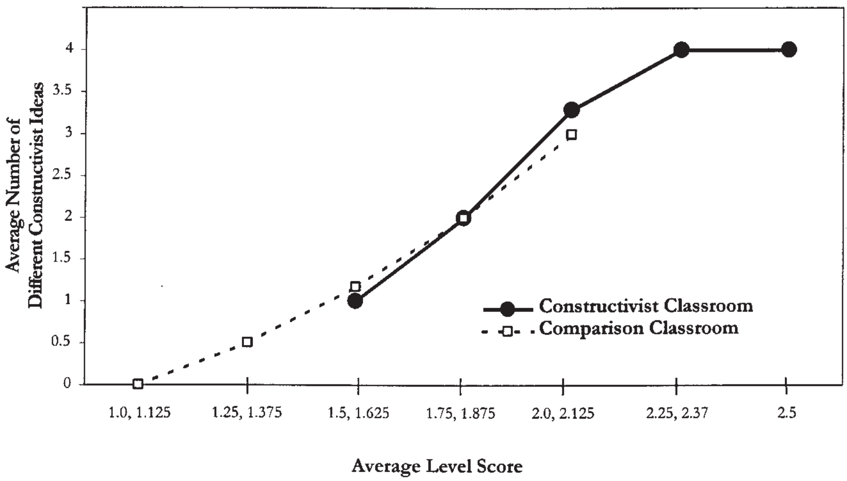

FIGURE 6 The relation between average level score on the nature of science interview and the number of different Level 2 ideas expressed for students in the two classrooms.

constructivist classroom mentioned all four of these ideas at least once across the four question clusters, and 83% of the students mentioned at least three of the four key ideas. Many students mentioned a given idea on multiple occasions as well.

Figure 7 shows some of the ways these four processes were elaborated and interrelated in students' responses as shown by the question cluster analysis. It also indicates how these processes interrelate with collaboration, a process students in the constructivist classroom highly valued. Supplementary analyses (presented in the next section) provide further support for the characterization of students' modal epistemology in Figure 7. One of these analyses provides a detailed description of how students view the role of social interaction in acquiring knowledge. The Discussion section provides a more extensive consideration of how all of the processes in this characterization interrelate.

Of course students in the constructivist classroom did mention simpler ideas in response to the four question clusters. For example, they referred to the gathering of information, mentioned journalistic questions, and talked about the testing of a simple prediction to find out if something works. We would not expect all mention of these ideas to drop out with a move to a more sophisticated epistemology. After all, scientists do gather information, and they do use experiments to answer specific questions or to test more concrete ideas (e.g., which medicine works better?). However, because these students had higher level notions that they discussed in their responses to these same questions, they can reframe their understanding of such activities in the context of a larger goal or enterprise.

Modal epistemology in the comparison classroom. Figure 8 diagrams the four key ideas that form the foundation for the Level 1 knowledge unproblematic epistemology of the students in the comparison classroom: (a) scientists do things or gather information (do things or gather information responses for Cluster 1), (b) scientists ask procedural or factual questions (procedural and journalistic questions for Cluster 2), (c) scientists do experiments to get products or to answer questions (try out/find cures and find answers responses for Cluster 3), and (d) information accumulates unproblematically (keep or abandon responses for Cluster 4). It shows how the ideas students have about the goals of science inform their ideas about the types of questions scientists ask, the purpose of experiments, and the nature of the change process. Fifty-two percent of the students mentioned all four of the ideas in Figure 8, and 89% of the students mentioned at least three out of the four ideas.

As evident in a supplementary analysis presented in the next section, students in the comparison classroom were much less concerned with issues of collaboration in knowledge acquisition than students in the constructivist classroom. Hence, collaboration is omitted from this diagram. Our analysis showed that the relatively few times these students did mention social interaction, however, it was primarily in relation to the Level 1 processes shown in Figure 8.

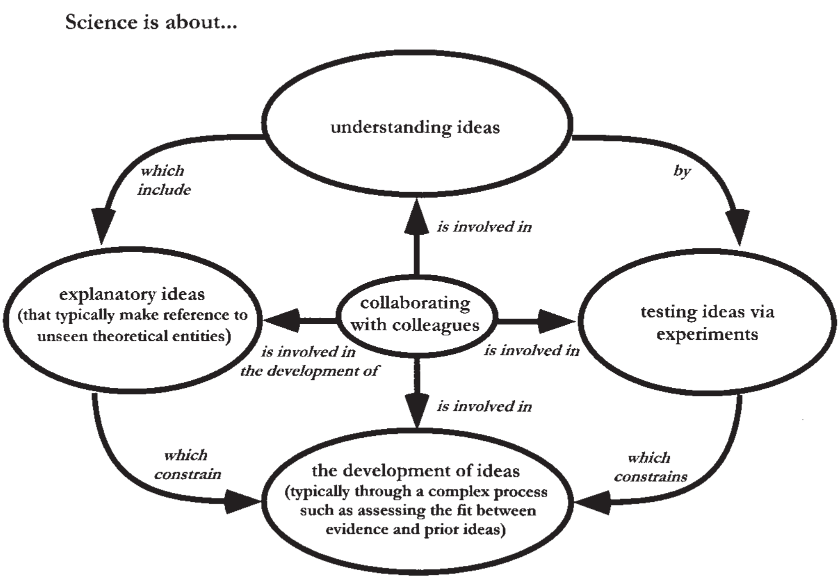

FIGURE 7 Network of ideas in the modal epistemology for students in the constructivist classroom. Labeled arrows indicate some (but not all) of the relations among these ideas.

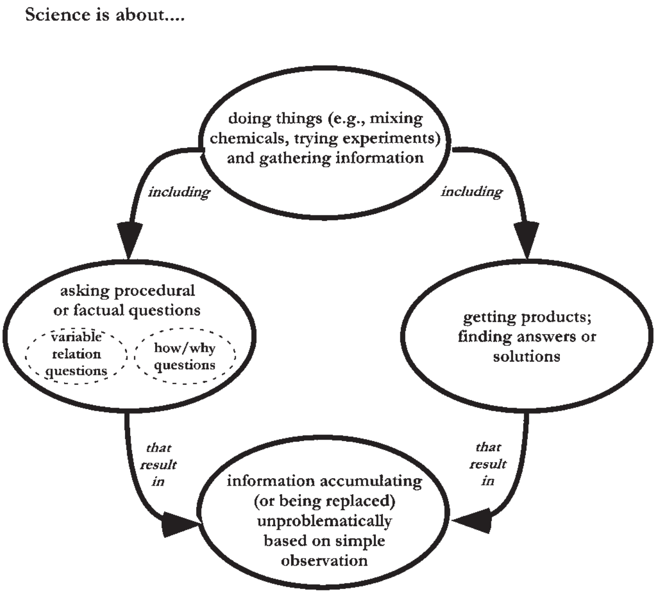

FIGURE 8 Network of ideas in the modal epistemology for students in the comparison classroom. Labeled arrows indicate some (but not all) of the relations among these ideas.

Although virtually all of the students mentioned these Level 1 ideas, only three students expressed such bare-bones views with no further elaboration. It was more common for students to have some Level 1.5 ideas or one isolated Level 2 idea in their responses. Most commonly, students with Level 1 scores went on to discuss either scientists engaging in thinking about data or ideas, asking a variable relation question, or asking a question about how things work or why things happen. Overwhelmingly, if students expressed just one Level 2 idea, it involved explanation. When students begin to consider these higher level ideas in isolation from other Level 2 ideas, their Level 1 knowledge unproblematic epistemology is not immediately undermined; however, their Level 1 views do start to enrich and become more complex.

Supplementary Analyses: Beginning Awareness That Knowledge Is Problematic

We carried out three supplementary analyses by reading the entire interview of students in both classrooms and looking for evidence of beginning awareness of ideas that would support the move to a knowledge problematic epistemology. In particular, we examined students' awareness of deep explanatory questions in science, complex criteria for evaluating ideas in science, and the importance of social interaction in the development of scientific knowledge. At issue was whether students in the constructivist classroom would show greater awareness of these ideas than students in the comparison classroom.

In addition, a fourth analysis examined students' responses to interview Questions 10 through 14—the portion of the interview that probes most directly for Level 3 rather than Level 2 constructivist understandings. These questions ask students to define the terms *hypothesis* and *theory* and to explain how scientists' theories affect the ideas they have about specific experiments. At issue was whether these questions would draw out any fully developed Level 3 insights among students in the constructivist classroom and, if not, whether these students were more likely than students in the comparison classroom to have some beginning understanding of these difficult issues.

Awareness of deep explanatory questions in science. One factor that supports a move to a knowledge problematic epistemology is an understanding that scientists often are concerned with "deep" explanatory questions: questions about unseen underlying causal mechanisms (and interactions between conjectured theoretical entities) rather than questions about direct causal relations between observable variables. Although establishing any kind of causal relation is complex and calls for interpreting a pattern of evidence, questions about unseen causal mechanisms call for even more complex and indirect arguments from evidence.

In this analysis, we first read students' entire interviews to identify all how and why questions (either explicit or implicit) that might be concerned with an underlying causal mechanism that involves a specific unseen theoretical entity. We then looked for evidence that students thought such questions were complex to answer, as further evidence that they were aware of the deep explanatory nature of these questions.

Significantly more students in the constructivist classroom (83%) generated potentially deep explanatory questions at some point in the interview than students in the comparison classroom (37%), $\chi^2(1, N = 45) = 9.36$, $p < .01$, two-tailed. For students in the constructivist classroom, these questions centered around four main topics: (a) How do atoms work? (b) What causes disease? (c) How do things move and what causes them to move that way? and (d) How do people learn and come up with their ideas? All called for serious theory building and were related to issues that the students had worked on in science class. A striking feature of these questions was that they contained either an explicit or implicit reference to ideas that concern theoretical terms in science; for example, ideas about atoms (including ideas about electrons, protons, and quarks), ideas about genes and DNA, ideas about gravity and other forces, and ideas about underlying thinking and learning processes.

We also examined the processes students thought were involved in answering such questions. We identified six processes (described in the next paragraphs) that all involved evaluation of ideas as part of a multistepped endeavor and we found awareness of these processes to be much more common among students in the constructivist than in the comparison classroom. This finding provided indirect support for our judgment that students in the constructivist classroom really were asking questions about some unseen mechanism or process, because it revealed that these students had begun to appreciate the complexity involved in gathering evidence for the workings of unseen causal mechanisms.

The first complex process, identifying a hypothesis and testing it, is the classic Level 2 process discussed earlier, in which one first comes up with a testable idea and then tests it. The second, making observations and inferences, is another Level 2 process in which one combines observation with knowledge or inferences to reach a conclusion, and data are seen as clues that must be pieced together to figure out an answer. For example, in discussing how one would investigate how atoms work, one student said, “Like sometimes they try to bounce atoms off of each other or slow them down or reflect shadows off of them to find out what they are like.” Another student discussed how scientists would figure out how kids learn: “By asking about different ways they have been taught and questions about how they have been taught and what they think … would be clues kind of towards how kids learn in a certain environment.” A third kind of Level 2 multistepped process focuses on the importance of analyzing or investigating ideas that come from different perspectives. For example, in considering how scientists answer questions about why something is moving, one student commented, “They’d take their ideas, sort them out, compare them to other scientists’ ideas and eventually come up with an answer they think is good for what they’re applying it to.”

The remaining multistepped processes were arguably more sophisticated (Level 2.5 in our system). For example, “making” answers is a process in which one explicitly recognizes that an answer is not out there to be found and then undertakes the challenge of constructing one. This seems to be a more sophisticated version of finding a hypothesis and testing it because it implies a recognition that hypotheses, as potential answers, are constructed. One student voiced this idea in the context of discussing why scientists who were concerned with researching questions about genes and DNA would not necessarily achieve their goals:

(S37): It depends on what kind of scientists they are. If they are like a nuclear physicist or something or always working on genes and DNA and altering different things, then I don’t think hardly any of them achieve their goals, ’cause it’s so hard … like the answer is not necessarily

> there—they have to like go out and *make* the answer—then I don't think they always achieve their goals.[5]

Other sophisticated multistepped processes occasionally mentioned included (a) striving to have ideas fit together and be consistent and (b) grounding a pursuit in a theoretical orientation (e.g., one would first have a theory about how something worked and that would tell him or her what to look for).

The majority of students in the constructivist classroom who asked clear explanatory questions (13 of 15) showed some awareness of the need for multistepped processes in answering these questions. Typically these students discussed one of the first three processes (making observations and inferences was the most common), although 1 student (S37) discussed almost all of them! Awareness of the complexity involved in answering explanatory questions was much less prevalent for the students in the comparison classroom (4 of 10).

Thus, 72% of students in the constructivist classroom, consistent with their move to a more constructivist epistemology and to a beginning awareness of knowledge problematic issues, were aware that scientists might ask a deep explanatory question and indicated that such a question would be complex to answer. In contrast, only 15% of the students in the comparison classroom showed this combined pattern, a significant difference, $\chi^2(1, N = 45) = 15.12$, $p < .001$, two-tailed.

Complex evaluation criteria. Another aspect of a move to a knowledge problematic epistemology is realizing that many scientific beliefs are not simply right or wrong, but, rather, need to be evaluated with respect to more complex criteria. In this analysis, we read students' entire interviews looking for appeals to more complex evaluation criteria. We also looked for explicit mention of knowledge problematic themes; for example, direct comments by students of the constructed nature of scientific knowledge.

Most students in both classrooms thought scientists were concerned with finding out if their ideas were right or wrong. The majority of students in the constructivist classroom (72%), however, also appealed to additional criteria for evaluating ideas during the interview. These included gauging whether ideas were useful, explanatory, made sense, fit with other ideas, or fit with a pattern of evidence. In contrast, only 2 students (7%) in the comparison classroom appealed to complex evaluation criteria: One mentioned the need for ideas to fit with each other and with a pattern of evidence; the other proposed the design of a controlled experiment.

In addition, 4 students from the constructivist classroom (S7, S19, S33, and S37) made explicit metastatements that science does not involve a simple process

---

[5]In student transcripts, ellipses indicate that some intervening words or comments by the student have been omitted. Questions raised by the interviewer are put in brackets.

of “finding answers” because scientific answers are just not there to be found. For 3 of these 4 students the metastatement came at the very beginning of the interview when they were asked explicitly about the nature of science and how scientists achieve their goals. For example:

> (S7): I mean you are not going to just *find* something. I mean you have to do research and stuff, and like, not discover it, but get an answer in your head. It’s not just there and you pick it up; it’s like you got to get it in your mind and then you got to try it out.
>
> (S33): OK. They don’t go out and find things. Like an idea is not, it’s not there. So they don’t go and get it. They have to like take their ideas and other people’s ideas and put ’em together, and then they’ll come up with a theory. And they, sometimes they try and make models and that might help them. And that’s it.

The fourth student’s metastatement (S37, cited earlier) came at the end of the interview when he was commenting on why scientists do not always achieve their goals.

Other explicit statements about difficulties of acquiring scientific knowledge came in response to interview questions about the nature of theories (to be discussed more fully in the final part of the Supplementary Analyses section). In those questions, 2 students made a distinction between facts and theories and discussed the conjectural nature of theories (S19 and S37), and 2 students talked about how theories can bias one’s interpretation of ambiguous experimental results (S33 and S37). The general theme—strongly held ideas are difficult to change and require multiple disconfirming experiments—was expressed by several other students in the constructivist classroom (S1, S7, S19, S28) at different points throughout the interview. In the words of one student (interviewer comments are enclosed in brackets):

> (S7): Because if your ideas are that strong you can’t just change them. You have to go at it a different way. [Why?] You would know you can’t go at it the same way cause it’s not going to do anything for you. But if you go at it a different way … that experiment might tell you something else, like it might be more obvious to you that you have to change your ideas. … I know that if I have a theory about something we are doing in science then I’m not going to change it just based on an experiment I did or something. I’m still going to want to do other things and go at it different ways to see if that is really true.

In sum, 72% of students in the constructivist classroom mentioned more sophisticated knowledge evaluation criteria at some point in the interview. These criteria (e.g., evaluating whether something is a better explanation or makes sense given one's prior ideas) provide further support for a constructivist Level 2 epistemology. Some of the students formulated an explicit distinction between finding and making answers or talked of the ways in which one's ideas can bias one's interpretation of results or be resistant to change (33%). In contrast, the majority of students in the comparison classroom (93%) talked throughout the interview of finding out if answers are right or wrong, an idea that is consistent with a knowledge unproblematic epistemology.

Ideas about social interaction. A third aspect of a move to a knowledge problematic epistemology is awareness of the diverse perspectives of individuals and the need to provide evidence and arguments that are convincing to others. In a third analysis, we read students' entire interviews to look for all instances in which they mentioned the role of other people in the knowledge acquisition process. The constructivist classroom featured a heavy emphasis on group work, exchange of views, classroom dialogue, and development of shared norms. Thus, we were interested in whether this emphasis would be reflected in the students' epistemological views. Would students in the constructivist classroom more frequently refer to social interactions as an inherent part of the knowledge acquisition process? If so, would they conceptualize these social interactions differently from students in the comparison classroom?

Two coders worked together to identify and underline all transcript portions that referred to social interaction and to develop a levels-based way of categorizing these social interactions.[6] Appendix C presents the details of this coding system. In brief, at Level 1, students talk of scientists as participating in concrete activities together or sharing information with each other. At Level 1.5, students talk of scientists as sharing, considering, and comparing ideas with each other. The purpose associated with these activities is to enlarge the scope of ideas considered, not to test, develop, or evaluate the viability of ideas. At Level 2, social activities go beyond simply sharing ideas to incorporate Level 2 notions of explanation, test, development, or understanding. For example, students talk of such activities as exchanging explanations about how things work, testing each other's ideas with experiments or with a social process, using other people's ideas to help develop one's own ideas, and trying to understand each other's ideas.

At Level 2.5, students either interrelate two or more Level 2 ideas in a social context (i.e., composite answers such as, they perform tests on each other's ideas

---

[6]After the coding system was developed, a third scorer categorized the ideas about social interaction that had been underlined to assess the reliability of the coding system. There was reasonable reliability about coding not only for the general level of a response (80% agreement using a four-category system of Levels 1, 1.5, 2, 2.5), but also for the specific subtype within that level (75% agreement on subtype that involved 12 subcategories across levels).

to understand each other's ideas), or they cite a more complex social process. These more complex social processes include gathering evidence to persuade someone else of the value of an idea or to convince someone that an idea is wrong, and working cooperatively with one's colleagues to build consensus while evaluating and developing ideas.

As expected, social interaction was mentioned more often by students in the constructivist classroom than in the comparison classroom. All students in the constructivist classroom spontaneously mentioned social interaction at least once during the interview. On average, they brought up social issues on 6 questions in the 23-question interview, with a range from 1 to 13 questions. Some students went as far as stating that social dialogue is necessary for success. In commenting on whether scientists always achieve their goals, one student asserted:

> (S35): No. … Like if other people didn't support their ideas, and other people didn't believe them, I don't think they could be successful on their idea. … Like I said: you need to hear what other people think to make your idea better.

In contrast, 30% of the students in the comparison classroom did not mention social interaction at all. Furthermore, the average number of questions on which social interaction was mentioned in the comparison classroom was only 1.1, with a range from 0 to 3 questions. This average number is much less than the average number for the constructivist classroom, a difference that is statistically significant ($t = 6.97$, $p < .0001$, two-tailed).

The two classrooms not only differed in how frequently social interaction was mentioned, but also in the ways students conceived of this interaction. Table 2 shows the percentage of students in each classroom giving each type of response. This analysis reveals that when students in the comparison classroom talked of social interaction, it was primarily in terms of Level 1 or 1.5 ideas. Level 1 ideas of sharing activities or information were most common, followed by Level 1.5 concepts of simple sharing of ideas. In contrast, the vast majority of students in the constructivist classroom expressed at least one Level 2 and one Level 2.5 idea about social interaction.

An analysis of the specific Level 2 and 2.5 responses shows the variety of ways that students in the constructivist classroom considered social interaction important in the knowledge acquisition process. Students did not merely comment on the importance of hearing the ideas of others. They generally went on to comment about how this activity contributed to the development of their ideas or the better understanding of their ideas. Sharing explanations with others was also commonly mentioned. The simple testing of each other's ideas with experiments or with a social process was not as frequently mentioned, but this was because the notion of testing was often articulated at a more complex level; that is, one needs to use evi-

TABLE 2
Percentage of Students Expressing Different Ideas About Social Interaction in the Two Classrooms

Level	*Response*	*Constructivist (%)*	*Comparison (%)*	*Significance Level*
1	Do activities together or exchange factual information	39	48	*ns* (p = .5509)
1.5	Exchange ideas with others	56	30	*ns* (p = .0855)
2	Develop ideas together	50	4	p = .0001
	Exchange explanatory ideas	44	4	p = .0005
	Understand each other's ideas	44	4	p = .0005
	Test each other's ideas with experiments	28	7	*ns* (p = .0672)
	Test each other's ideas with social process	17	0	p = .0282
	At least one Level 2 response	94	19	p < .0001
2.5	Use evidence to establish the viability of an idea to others	56	4	p < .0001
	Composite response	28	0	p = .0030
	Operate with social principles	22	0	p = .0095
	Build consensus about ideas	22	0	p = .0095
	Influence each other's interpretations	6	0	*ns* (p = .2247)
	At least one Level 2.5 response	78%	4	p < .0001

dence to persuade someone else of the value of one's idea or to convince someone else that his or her idea is wrong (a Level 2.5 idea).

Using evidence to establish the viability of an idea to others was, in fact, mentioned by more than half of the students in the constructivist classroom. Other ideas, each mentioned by about one quarter of the students, were combinations of two or more Level 2 concepts in a social context, comments about upholding the social principle of testing ideas prior to admitting them into the social arena, and finally, comments about working cooperatively with colleagues to build consensus on ideas and to strive toward fitting a new idea with others' preexisting ideas. Less frequent, but similarly interesting, were comments about working to influence each other's interpretations.

In summary, students in the constructivist classroom mentioned ideas about social interaction both more frequently and in more sophisticated ways than students in the comparison classroom. Their comments depict scientists putting substantial time, care, and effort into the challenges of persuading each other, meeting social standards, moving toward consensus and influencing each other's interpretations while they strive to explain phenomena and to develop, evaluate, understand, and clarify ideas.

Conceptions of hypotheses and theories. In the middle of the interview, we asked students whether they had heard the words *hypothesis* and *theory* and, if so, what they thought hypotheses and theories were. At issue was whether students think of hypotheses and theories as equivalent to procedural directions or facts (i.e., how to do things or what happens) or as equivalent to explanations for how things work. Also at issue was whether students differentiate between hypotheses and theories and whether they recognize that theories constrain specific hypotheses.

All students in the comparison classroom had heard the word *hypothesis*. Indeed, almost all used exactly the same words in defining a hypothesis—it was an "educated guess"—which suggested that they had been given an explicit definition in science class. Follow-up probing indicated that, for most students, the educated guess was about what materials to use in an experiment (what things would work best) or what would happen while following one's procedure. Only one student made some reference to explanation; she defined a hypothesis as "a logical explanation or guess."

A similar picture emerged for their understanding of the term *theory*. Although students were more idiosyncratic in their responses to what a theory was, their responses were fairly low-level or ambiguous. Some were unfamiliar with the term or defined it fairly concretely in terms of the steps used to do something, a way of saying something, a topic, what happens, or a guess. Others initially mentioned that it was an idea about something or "something that you think," but then went on to say that it was an idea about what to do or what happens. Still others left the nature of the idea unspecified. Only one mentioned that it was an idea about "why things happened" and another described it as a "viewpoint."

Unlike the students in the comparison classroom, no student in the constructivist classroom was familiar with the word *hypothesis*. (Their teacher later informed us that this term had never been explicitly introduced in their classroom.) They all, however, were familiar with the word *theory* and their responses were generally either ambiguous or at a higher level than those of students in the comparison classroom. Only two students in the constructivist classroom described a theory specifically as an idea about what happens or how to solve a problem. Many (55%) described a theory ambiguously as a person's idea about something, and left the nature of the idea unspecified. The remainder of the students (33%) specifically stated or implied from their examples that theories are concerned with ideas about how things work.

Two students in the constructivist classroom struggled at length to discuss the conjectural nature of theories and how they are evaluated with indirect arguments or evidence. For example, one of these students discussed how a scientist's theory about atoms was like one's attempts to theorize about the contents of a box, where one cannot look inside the box or directly see the contents:

(S37): A theory is maybe a thought of what you think it is. … There's something in a box and you couldn't see it, and you knew it made a certain noise when you shook it and then it was heavy, you'd have, you could develop a theory, well I think it's a rock or a weight or something, because you are not sure, but it's like what you think. [So do you think scientists have theories?] Scientists? I think they do have theories. Because, let's see, there's Helmut Fischler who came to our class, and he works on the atom and all he has is theories. I mean, cause you can't really see an atom, so he has to think of, well "this is what makes sense," like this part of the atom needs something to hold it together, so let's see, one part of the atom would need another part to help it or something. His theory would be he thinks there are protons and neutrons of an atom, because it makes sense to him, although for most people it's almost a fact that there are those things. He thinks electrons go around in little circles, because it makes sense to him. Nobody really knows the truth, but it's his theory because that's what makes sense to him.

The other student also discussed the role of indirect argument and evidence and the use of analogies in the development of theories about how atoms work:

(S19): A theory is an idea that you have … . You can't have a theory about a chair—that the chair is there—because that's a fact, so theories can't be facts. They would be just ideas that you may be able to prove with some experiments or you may not be able to prove. [Do you think scientists have theories?] Yeah. I think that is what their work is based on. [In what way?] Like people who work with atoms … There wouldn't be even a field or their job if there wasn't a person who had a theory of the atom, and some people, their job is like just thinking up of theories about atoms, good explanations. [And other people, what's their job?] To try experiments out about the atoms. Like sometimes they try to bounce atoms off each other or slow them down or reflect shadows off of them to find out what they are like. [Did you have someone come talk to you about this?]. Yeah, some people came and we were talking about our analogies of what they do. [Analogies, what do you mean by analogies?] One analogy was like what if the science room was a dark room … . There was something in there, but we didn't know what it was. And all we had was an open window, but there was no light shining in … . And then we were thinking of bouncing the tennis balls off and finding what sound it made, or if anything flew off of it, if it was fragile or something concrete. And *that*, we sort of figured out *that*, was an analogy for—bouncing the tennis balls were little atoms, and in the room was a big atom and we were bouncing the little atoms off the big atom to find what form or what the pieces were of the atom, because there was no other way we could see it.

Thus, across students in the two classrooms, there was a range of meanings given to the word *theory,* progressing from a fairly restricted notion of theory as involving something concrete (i.e., what to do, what to say, etc.) or some concrete idea (i.e., an idea about what to do or what happens), to a broader notion of theory as involving any idea you have (including both more concrete and more explanatory ideas), to theories as specifically concerned with conjectural and explanatory ideas. In general, awareness of the explanatory nature of theories lags behind awareness that scientists ask explanatory questions. Some students who gave examples of explanatory questions still seem to have a broader meaning for theory that includes any idea that scientists have.

Question 14 in our interview (Do you think a scientist's theory influences his or her ideas about specific experiments?) was very difficult for students in both classrooms. It was designed to probe for an awareness that scientists' theories constrain their generation of specific hypotheses and interpretation of experimental results (Level 3 understandings). In fact, no student in either classroom interpreted the question in both of these sophisticated ways. Some students explicitly balked at the presence of the words *theories* and *ideas* in the same question—arguing that they were the same thing so that the question did not make sense. Others seemed to ignore the second mention of *ideas* and simplified the question to "How do scientists' theories (or ideas) affect their specific experiments?" or reversed the question to answer "How do scientists' experiments affect their ideas?" Still others thought theories affected experiments by specifying what to do or what to try out. In these responses students often discussed getting a good outcome or getting something to work. There was surprisingly little focus on designing experiments to evaluate ideas where the ideas were something more complex than an idea about what works. Finally, a few students could articulate no relation at all between a theory and experiment.

Seven students (39%) in the constructivist classroom, however, did attempt an interpretation of the question that had an explicit role for both *theories* and *ideas* and that went beyond the idea that theories were simply the ideas tested in experiments. One of these students said that a scientist's theory about how atoms work provides reasons for his or her ideas about atoms, but he did not elaborate further. The others talked of theories as affecting the scientists' perception or understanding of experimental results. The 2 students with the most sophisticated responses (S33 and S37) discussed how the scientists' commitment to their ideas could lead to bias in interpreting ambiguous perceptual information; hence it would take scientists a number of trials to overcome this bias (these students were lab partners who encountered this problem in their investigations of the motions of objects). Others talked more simply of theories affecting what events one chooses to focus on or look at in an experiment, or whether one will initially believe or understand one's results.

In summary, these explicit questions about hypotheses and theories were difficult for students in both classrooms and failed to elicit any Level 3 understandings.

Nonetheless, students in the constructivist classroom showed more insight on these questions than students in the comparison classroom in that (a) one third of them defined theories as specifically concerned with explanatory ideas, and (b) more than one third understood that theories could influence other mental processes (other ideas, what one chooses to look at, how one interprets or evaluates what one sees, etc.). In contrast, although all students in the comparison classroom knew the word *hypothesis*, their definition of "educated guess" (either about how to do something or what will happen) was still consistent with a knowledge unproblematic epistemology. Furthermore, they generally thought of theories as activity oriented—as directly affecting what one chooses to do rather than how one thinks.

DISCUSSION

Can Sixth-Grade Students Develop a Sophisticated, Constructivist Epistemology of Science?

The main question addressed by our study was whether sixth-grade students would be able to develop a sophisticated constructivist epistemology of science by participating in an elementary science curriculum that supported the development of such an epistemology. Our data suggest that they can. In particular, students in the constructivist classroom were centrally aware that science involved the development and modification of ideas about how the world works, that these ideas take work to develop and understand, that experiments are useful both as a means of clarifying and testing ideas, and that collaboration is important in all aspects of the process. Furthermore, the understandings revealed through our analyses of the Nature of Science Interview data were consistent with prior data gathered by Hennessey for these same students in the natural context of her classroom. Significantly, these students' understandings about science go well beyond what has been previously reported in the epistemological literature for students of this age, and therefore provide further evidence against the view that there are biologically based developmental constraints on young children's thinking of the type envisioned by Piaget. Although their understandings fall short of a Level 3 epistemological stance, these students have made progress in appreciating some of the kinds of mental and social work that are part of the process of scientific knowledge acquisition. Let us now consider each of these points in greater detail.

Core epistemological ideas about science for students in the constructivist classroom (based on the Nature of Science Interview data). Figure 7 diagrams the network of Level 2 ideas that we identified for students in the constructivist classroom based on their responses in the Nature of Science Interview. It also shows some of the ways these ideas were beginning to be interrelated in students' responses.

The facts that the students consistently referred to a variety of Level 2 notions throughout the interview, that they had begun to interrelate these notions explicitly, and that each of these notions supports the differentiation of ideas from evidence suggest that these students had developed a qualitatively different way of thinking about science from students in the comparison classroom. These Level 2 ideas also were supported by the students' understanding of the importance of collaboration in scientific inquiry: Collaboration was seen as integral to the processes of understanding one's ideas and sorting out their status, generating explanations, and testing and developing ideas.

At the top of the diagram is concern with understanding one's ideas about the world. Almost all of these students expressed this idea at some point in the interview, and for most students it was an idea mentioned in the first cluster of questions. Thus, it seemed to be a framing idea for them (which is why it is placed at the top of the network). Furthermore, they all seemed to realize that understanding ideas takes work. They viewed discussion with others and active experimentation as both contributing to this process.

Related to the concern with understanding ideas is the awareness that scientists are concerned with explanation and that different people have different ideas about how to explain events. Scientists ask many types of questions, not just factual questions, such as, "Is this apple red or green?" Rather, they are concerned with understanding how things work and why they happen. In seeking explanations, they are concerned with theoretical concepts (such as atoms or forces), and they ask questions about these entities (e.g., What are atoms like? What forces are at work in different situations?). We speculate that it is because the students themselves had been concerned with investigating deeper explanatory ideas and with considering the often initially unintelligible explanatory ideas of their peers that they had come to realize that they have to "work" to understand this type of idea.

Closely related to a concern with understanding one's ideas and striving for an explanation are the concerns with testing and developing ideas through experiments. If the scientist is not simply looking for observable results (i.e., did it work or not), then it follows that experiments do not provide answers but, rather, provide evidence for or against an underlying idea. Scientists engage in trying to prove ideas right or wrong to each other. To be convincing (to oneself and others), one often needs to test an underlying idea in different ways with multiple experiments. Ideas are complex and have multiple parts that get changed, rearranged, and revised over time. Furthermore, doing multiple experiments is intimately involved with the process of developing an understanding of ideas in a community of inquirers. In such a community, one's ideas not only have to fit with the evidence, but also with one's own ideas and the ideas of others. Thus, the overarching goal of understanding one's ideas, which is central in the constructivist classroom, leads students to engage in and recognize the importance of the complex and sustained kinds of work that are needed for these understandings to evolve.

Relation to previous data gathered in the constructivist classroom. Further evidence that students in the constructivist classroom genuinely understood that science (and science learning) involves the active construction and modification of ideas, rather than the simple accumulation of facts and information, comes from an analysis of data gathered earlier in the year by Hennessey in the natural context of her classroom. In that study, these same sixth-grade students had been asked to respond to one of two questions about their thoughts: One question asked them to describe their views of what science is, and the other asked them to describe their views of the nature of learning in their science classroom. Students worked individually during class time for about a week to write essays expressing their individual views. A little over half of the students chose to write about the nature of science, and the others wrote about the nature of learning in science class.

Among the essays written about the nature of science, all students commented on the central role of ideas in science and the importance of working to understand ideas—both framing themes expressed in responses to our Nature of Science Interview as well. In addition, all students explicitly commented on the ways scientists' ideas guide their inquiry processes. Some students commented that scientists must use their ideas, along with the ideas of others, in developing their theories; other students commented that scientists must use their ideas in planning experiments, interpreting experimental results, or, more generally, in guiding everything they do. The guiding role of ideas had been expressed in their Nature of Science Interviews, although not always as explicitly or eloquently as in their classroom essays. The following extracts from essays of four different students serve as examples (Hennessey, 1994a).

(S15): Science is many different things to me. It's about understanding what I think and why I think it. It's about understanding how things work. For example, physicists want to know how atoms were formed and they gather together to create matter. However, physicists cannot really know for sure "how" the atoms were formed because the universe was not in existence at the time they were created. So physicists develop a theory of how they think atoms were created. This theory is more than just a guess, it is based on their own ideas of what the universe could have been like at the very beginning. As their ideas about the early universe change so does their theory of how the atoms were created.

(S30): Science is also about ideas and understanding. You would have to understand what you're dealing with when you're trying to find answers to questions. When you are finding answers, however, your own ideas about the question guide your plans for the experiments that you do to answer the questions. When you are trying to find the answers to your questions, your own ideas become part of the answer. You use your ideas and past experiences to guide everything you do.

(S23): Dealing with your ideas is the most important thing that you can do. Your ideas are what you use to interpret your experience. Some ideas are more developed than others and you can use these ideas to plan science experiments or interpret the results of the experiments. Other ideas are fuzzy or not very clear. When ideas are not clear they cannot be used to explain or interpret anything.

(S37): I think science changes because people's ideas change over time. This even happens in school science. For example, when someone tells you their ideas you may or may not understand it. However, if they change their explanation a little then you can understand it. Or when different people in a class explain their thinking about something we are all working on, soon different people in class begin to change their thinking and so do I. That's how I develop my ideas. I discuss with other students and I listen to their explanations. I try to see things from their perspectives and they try to see things from mine. All of us begin to develop ideas that are a combination of what we hear or discuss—that's how I change my thinking. I think people who are scientists do the same thing. Only when they change their ideas or describe them from a different perspective then science itself changes.

In the essays about the nature of learning in Hennessey's science class, once again all students focused on learning as a constructivist process of trying to understand and develop one's ideas. They also all commented on the central role of collaboration and interpersonal exchange in this learning process: Sharing ideas with others not only helped them to understand their own ideas, but also provided a valuable source of new ideas. These themes had been expressed widely in responses to our Nature of Science Interview as well. The one theme that was more centrally and explicitly articulated in their classroom essays than in the Nature of Science Interview was the idea of perspective taking. Significantly, the majority talked of learning as involving active perspective taking, considering and comparing multiple points of view, and figuring out how perspectives relate to each other. For example, the following extracts illustrate the ways students described the perspective-taking process in their essays (Hennessey, 1994a).

(S39): Understanding my ideas and how they relate to other people's ideas is not always easy and takes a lot of time. The people you are working with in your lab groups are also trying to explain their ideas and all of us are trying to explain our ideas about the experts' ideas. This can get to be confusing because different people see the same thing from different perspectives. It's nice to know that there is no one way of looking at things like Mr. Newton's ideas about gravity or Mr. Einstein's ideas about gravity—they are so different. I guess that is because they worked from different perspectives.

(S7): Learning is considering my own ideas and how they fit together. I like to consider my own thinking because then I can understand more about why I think the way I do. I can also learn about the science communities' ideas the same way. For example, I can compare my classmates' ideas to mine and by doing this it helps me build a bigger and stronger or more complex description of what we are talking about. My classmates and I also can do the same thing when we are considering a scientist's ideas. It's really the same thing you know—thinking about someone else's ideas and trying to see it from their perspective ….

To learn I think you need a good understanding of your ideas about the topic before you can do anything else. Like when I stopped to think about my ideas of how gravity worked, they made perfect sense to me. I remember the big debate we got into in third grade over how gravity worked—did it push or did it pull! Now the debate is much more complex and centers around comparing and contrasting our ideas to Isaac Newton's ideas. And trying to figure out what he meant by parts of his theory; like the influence between two objects, no matter how far apart, never decreases to zero—now that I find intelligible but certainly not plausible!

(S33): I think the most important part of learning is being ready to change the way I look at things. Looking at things from different perspectives helps to make it easier for me to think about the things better. It makes it easier to change my ideas, to find different ideas more plausible than mine, and explain complicated ideas better. Say you didn't have any ideas that you thought about on a part of science and you looked at a researcher's work. I don't think it would make any sense to you at all because you didn't even have one thought of your own about the researcher's work. You couldn't use your thoughts to help understand the researcher's work. But, if you had thoughts of your own, you could use them to understand the ideas of the researcher's work and it would be much easier to explain the researcher's work to someone else.

(S45): Learning is easier to do when you compare and contrast your ideas with someone else's ideas. Comparing and contrasting helps you think of things from a different perspective than your own. When you think of things from different perspectives then your ideas began to mix with the other perspectives and finally your ideas change over time.

In these essays on learning, students also spontaneously distinguished between the active process of learning they experienced in Hennessey's class (a process that involved understanding ideas, fitting them together, relating them to other people's ideas, and applying them in new contexts) and more passive processes that they had experienced in other subjects (memorization and repeating facts without understanding). These comments provide one piece of evidence that these students were developing different epistemological stances in different domains that reflected their contrasting educational experiences. Even more direct evidence on this point came from a whole-class discussion these students had with a science educator, Peter Hewson, and a physics educator, Helmut Fischler, on these very issues during a visit to their sixth-grade classroom (M. G. Hennessey, personal communication, August 1999). In that discussion, the students expressed disbelief that subjects such as math or reading could be taught in the more constructivist fashion they had experienced for science, and they challenged Hewson and Fischler to explain how it could be done!

Relation to previous findings in the epistemological literature. With respect to the epistemological literature previously reviewed, the kinds of constructivist insights that were expressed by students in Hennessey's classroom, in both their responses to our Nature of Science Interview and their classroom essays, go well beyond the kinds of insights that would be expected for students of this age.

The prior literature on students' epistemologies of science using Carey's Nature of Science Interview showed that middle school and even high school students with traditional schooling backgrounds typically responded in simple Level 1 fashion. The percentage of students that expressed consistent Level 2 scores was small: 0% (Carey et al., 1989) and 3% (Honda, 1994) for two samples of 7th graders and 25% for a sample of 11th graders (Honda, 1994). The low rate of consistent Level 2 responding among the current sample of sixth graders in the comparison classroom (4%) fits with this prior literature. Average level scores for the previous studies generally increased after exposure to an innovative curricular unit about the Nature of Science designed to teach Level 2 and 3 points. The only sample for which a majority of students achieved average Level 2 views, however, was a sample of 11th graders: 64% of 11th graders (Honda, 1994) compared with only 18% (Carey et al., 1989) and 30% (Honda, 1994) of 7th graders, in two separate studies. Although Honda's criteria for Level 3 responding was less stringent than ours, Level 3 responding was quite rare in her studies and no student achieved consistent Level 3 responding.

In light of this literature, the achievements of the students in the constructivist classroom are particularly noteworthy: 83% of the sixth graders in the constructivist classroom had an average level score of at least Level 2, and all of these students had average level scores that were greater than 1.5. These findings show that the differentiation of explanatory ideas and evidence need not be limited to just a few precocious sixth graders, but is well within the grasp of an entire classroom of students. Indeed, more 6th-grade students expressed consistent Level 2 views than the 11th-grade students in Honda's (1994) study who participated in an innovative Nature of Science unit. Furthermore, the Level 2 epistemology achieved by students in the constructivist classroom was more elaborated and constructivist in tone than the Level 2 epistemology described by Honda.[7] In that coding system, the hallmark of Level 2 is an awareness that scientists do experiments to test their ideas. Hypotheses, however, are essentially descriptive rather than explanatory. Thus, for Honda, Level 2 is a simple inductivist epistemology. In contrast, students in the constructivist classroom mentioned not only the notion of testing ideas, but also the notions of explanation, development, understanding, and collaboration. They did not think that scientists got their ideas just by looking at experimental evidence. They were aware of the mental and social work it takes to understand and develop one's ideas and of the fact that explanations often concern unseen theoretical entities. Each of these additional ideas not only contributes to these students' differentiation of explanatory ideas and evidence, but also gives their epistemology a decidedly constructivist tone.

The constructivist insights achieved by Hennessey's students also go well beyond the new insights of elementary school students that are commonly reported in the theory of mind literature. For example, studies have shown that elementary schoolchildren come to understand (a) that simple exposure to an information source is not enough to know about something if one has not attended to the information or if the quality of that source is impaired (Montgomery, 1992; Taylor, 1988), and (b) that one can know something even if one has not directly seen it if one can logically infer it from the information given (Sodian & Wimmer, 1987). Wellman (1990), however, distinguished between becoming aware of the mind as an active agent and having a sophisticated constructivist theory of knowledge and truth. Whereas elementary schoolchildren become aware that prior knowledge may influence the accuracy of one's perception of events, they do not yet see prior knowledge as essential to the knowledge construction process or as directing its course. They also do not think of truth as elusive or relative to one's framework of inquiry.

In light of this literature, it is significant that the sixth-grade students in the constructivist classroom understand that scientific knowledge grows out of and depends on the prior ideas they and others hold, and that individuals have different starting ideas that influence their sense-making efforts. In their responses to the Nature of Science Interview, these students clearly recognized that scientists have initial ideas that are subject to evaluation, revision, elaboration, and development. They were also deeply aware of how knowledge-building efforts are enhanced by collaboration and consensus building, understandings that have not been discussed much in the prior epistemological literature and which Driver et al. (1996) found absent from the simple inductivist epistemology embraced by high school students. Furthermore, in their classroom essays, the students in Hennessey's classroom were able to reflect on these starting ideas as perspectives; to identify and

[7]An important feature of Honda's (1994) work was the effort to clarify the distinction between Level 2 and Level 3 understandings. We note that in our current system, what Honda called Level 2 would be considered Level 1.5 and what Honda called Level 3 would be considered Level 2.5.

bracket the perspectives of self, other, and the science community; and to be actively engaged in thinking about how these different perspectives relate to each other. They also made a distinction between the cognitive activities of memorizing and understanding, which Fabricius, Schwanenflugel, Kyllomen, Barclay, and Denton (1989) found to be a much later development. Students noted that they could repeat and remember things that they have not fully understood and argued that true learning involves understanding, not just memorizing. In all of these ways, the students in Hennessey's classroom have developed a more sophisticated, constructivist epistemology than has been previously reported for students of this age or than would be expected by those espousing the Piagetian hypothesis that elementary schoolchildren are "concrete" thinkers.

Has this greater awareness of the role of ideas in knowledge construction undermined a belief in absolute truth and led these students to hold a radical relativist view in which all viewpoints are equally valid and conflicts in opinion cannot be resolved? Or, were they able to develop a commitment to evaluating different points of view in light of arguments and evidence and an openness to changing their views? As Hofer and Pintrich (1997) noted, "Openness to new interpretations is a key element of King and Kitchener's (1994) highest stage of reflective judgment and D. Kuhn (1991) speaks of evaluative epistemologists (the highest level) as open to the possibility that their theories may be modified by genuine interchange" (p. 120). Unfortunately, our Nature of Science Interview did not pose direct questions about the certainty of scientific knowledge or about how scientists would resolve conflicting claims, questions that were used in past research to elicit radical relativist views. Hence, we do not know what these students would say to such direct probes. What we can say, however, is that radical relativist views were not expressed in their responses to our Nature of Science Interview. Instead, their responses throughout the interview showed that these students saw scientists as committed to evaluating their ideas against multiple standards: the evidence from one or more experiments, the scrutiny of peers, and their own individual sense-making efforts. Furthermore, in their reflections on their own classroom learning in their classroom essays, it is striking how much value the students placed on considering the perspectives of others and being open to changing their own views. Thus, some important intellectual attributes and dispositions, commonly attributed to highly mature learners, already appeared to be present, at least in nascent form, in these sixth graders.

Limitations in students' epistemological views. Although the students in the constructivist classroom had developed and elaborated a rich Level 2 epistemology, their epistemology fell short of being the Level 3 knowledge problematic epistemology described by Carey and Smith (1993), in which students make a principled distinction between framework theories and more specific hypotheses, and in which they have a detailed grasp of the logic of hypothesis testing. Given their ability to develop an epistemological stance toward science that acknowledges the central role of ideas in knowledge acquisition, such limitations may reflect limitations in domain-specific knowledge of scientific theories and methodologies rather than limitations in a general capacity to reflect on their ideas. Developing a Level 3 epistemology of science certainly would require experiences with more sophisticated theory building using careful scientific methodology than these students had.

In what ways do students' ideas fall short of these Level 3 understandings? First, students in the constructivist classroom had not yet defined scientific theories as a coherent set of principles or concepts that are used to explain a wide range of phenomena and that constrain the generation of specific hypotheses; nor did they explicitly see the goal of science as developing increasingly more adequate explanatory theories. Rather, they talked more simply of theories as "their ideas" or at best "their ideas about how things worked," and they talked of the goals of science as "understanding their ideas," "developing their ideas," or trying to understand how something works or why it happened.

Second, in giving examples of scientific questions, students did not give examples of multiple levels of chained questions that they saw as intrinsically interrelated. For example, the process of answering a broad theoretical question includes giving operational definitions for key terms and asking more specific questions about relations among measurable variables. Students in the constructivist classroom generally gave examples of broad theoretical questions, but they did not explicitly comment on the subquestions involved in making operational definitions or in providing strong evidence for a causal relation.

Third, students in the constructivist classroom did not talk of experimentation as the controlled manipulation of environmental conditions that enables scientists to empirically distinguish alternative causal hypotheses. Nor did they discuss hypothesis testing as a means of providing indirect evidence for or against a larger theory. Rather, they talked more simply of an experiment as a means of testing, clarifying, or developing one's ideas and made no reference either to control groups or to a distinction between correlational and experimental designs.

Finally, in discussing the processes by which ideas change in science, students did not state explicitly that theories constrain the generation of new ideas. They also did not make an explicit distinction between normal science (in which one adds incrementally to an existing theory) and revolutionary science (in which a new theory ascends above an older theory). Rather, they talked more simply of change as involving the development of ideas, of new ideas as needing to make sense or fit a pattern of data, or of difficulties related to changing a whole theory either because it had multiple parts or because it was something that was strongly believed.

Our data from the Nature of Science Interview do, however, highlight the ways that these students' Level 2 ideas were preparing them to understand Level 3 issues, and the ways that they had already begun to appreciate the problematic nature of scientific knowledge. Students in the constructivist classroom did understand that scientists' ideas play a guiding role in inquiry—experiments are designed to test scientists' ideas and the goal of science is to gain a better understanding of their ideas. Many were aware that ideas had multiple parts that had to be pieced together and, hence, that a series of experiments was needed to build a better understanding of these ideas. The supplementary analyses revealed that most of these students were implicitly aware of the conjectural nature of scientists' ideas (as demonstrated by their examples of scientists' ideas and the fact that they realized that some multistepped process was needed to evaluate these ideas) and of the need for these ideas to be coherent (by talking of the need for ideas to "make sense"). A few went beyond these implicit understandings with explicit comments that scientists make rather than find ideas and that ideas need to fit together and be consistent. A number of students noted that scientists' ideas can affect what they choose to look at, bias their interpretation of results, and resist change. By expressing that coming to know something in science requires explaining phenomena using conjecture as well as developing, evaluating, understanding, and clarifying conjectural ideas, students in the constructivist classroom showed an awareness that scientific knowledge is problematic in nature. In their awareness that coming to know something in science requires scientists to contend with the perceptions and reasoning of social peers, students in the constructivist classroom showed an even richer sense of the problematic nature of scientific knowledge. That is, they showed some appreciation of how the fuzzy factors of human perception and reasoning constrain observation, conceptualization, and interpretation of phenomena, and how the fuzzy factor of human judgment constrains decision making and consensus building regarding which ideas are viable and which evidence is admissible.

Implications for the Teaching of Elementary School Science

Our work demonstrates that school science experiences can dramatically affect the development of epistemological thinking about science during the elementary school years. More specifically, the sixth-grade students in the constructivist classroom had clearly developed a more constructivist epistemology of science than students in the comparison classroom (or, for that matter, the students in previous studies). We would argue that the main factor responsible for the two groups' different epistemological stances toward science was the difference in their elementary school science experiences. Both groups were the same age (to control for maturationally based developmental factors) and demographically quite similar (to control for the influence of parents and outside-of-school experiences with science). Furthermore, other school subjects (for students in the constructivist science classroom) were taught from more traditional epistemological perspectives, making it unlikely that the students developed their constructivist insights from these other elementary school experiences. Both groups even had similar amounts of elementary classroom time spent on science. The main difference between the two groups was the target epistemology of science that the teachers (implicitly or explicitly) aimed to help their students develop and applied when designing their science curricula.

Other researchers (e.g., Brown & Campione, 1994; Lehrer et al., 2000; Metz, 2000; White, 1993) have already demonstrated that elementary schoolchildren are more "ready" to engage with issues of theory building and data gathering than has been assumed by those operating from a Piagetian-based developmental constraints perspective. What the research reported here adds is evidence that they also are able to build more sophisticated epistemological understandings about science than has been assumed or demonstrated in the prior literature. We attribute these developments to the innovative educational environments that provide intensive teacher scaffolding and support for student inquiry. Taken together, both kinds of research not only challenge the prior conventional wisdom about the kinds of intellectual demands a developmentally appropriate elementary science curriculum can and should make, but also begin to provide valuable alternative models of what a truly empowering and effective elementary school science curriculum can and should be like.

Design Features of an Elementary Science Curriculum That Appear to Support the Development of a Constructivist Epistemology

But what was it about Hennessey, her teaching, and her "constructivist" classroom environment that contributed to the tremendous growth in her students' epistemological views? We believe that many coordinated aspects of her teaching approach were essential. Although our study was not designed to address this issue directly, we conclude by discussing what some of these coordinated features may have been and the ways they may have supported epistemological development. Our discussion is informed by what we know about Hennessey's classroom from the work of those who have studied her classroom directly (Beeth, 1998; Beeth & Hewson, 1999a, 1999b; Hennessey, 1994b, in press; Hennessey & Beeth, 1993) and from our own informal observations. They also are informed by prior research and theoretical writings on features of classroom environments that may be particularly important in promoting epistemological understanding and conceptual change.[8]

[8]See Lederman (1992) for a review of studies of classroom environment features (at the high school and college level) that are associated with growth in student understanding of the nature of science. The findings from those studies are consistent with the arguments made here about critical features of an elementary school science classroom.

Authentic inquiry. First, Hennessey gives her students responsibility for managing most aspects of their own inquiry. Although she generally begins units by giving students a set of phenomena to explore, it is then the students' task to record the questions they have about these phenomena, to select questions they want to pursue further, and to plan ways to pursue their investigations. In these respects, Hennessey routinely involves her students in what Roth and Roychoudhury (1993) called *authentic contexts* for scientific inquiry: contexts in which students have responsibility for posing questions, generating methods, and analyzing data. This form of more "open" inquiry starkly contrasts with the more "closed" laboratory exercises used in traditional science classrooms, in which the problems, methods, and often even expected answers are given to students ahead of time. Some advantages of more authentic inquiry are that it ensures that the questions investigated make sense to the students (as they are grounded in what they know and think) and it elevates student motivation, interest, and involvement. In addition, as Chinn and Malhotra (in press) pointed out, there is an even deeper reason that involving students in more authentic inquiry may be particularly important in promoting epistemological development. By leaving open both the question of problem and method, students must confront a number of thorny issues about the interplay between theory and evidence that never get raised in more standard exercises. These authors analyzed the detailed ways in which more authentic experimentation and inquiry supports the development of a more constructivist epistemology and traditional laboratory exercises support a more inductivist or positivist view. Significantly, many of the innovative curriculum units developed for elementary school students in the last decade have all involved students in authentic inquiry (Brown & Campione, 1994; Lehrer et al., 2000; Metz, 2000; Roth, 1996; White, 1993).

Generative problems. Second, Hennessey selects initial problems that invite her students to consider issues of deep disciplinary significance. Gardner, Perkins, Wiske, and colleagues called these *generative topics* (Gardner, 1999; Wiske, 1997): topics that open up rich veins of inquiry within a discipline, such as science, mathematics, history, or the arts. For example, in one curricular unit, Hennessey's students explore their ideas about the day–night cycle and the causes of the seasons in ways that encourage them both to think about the relations between the earth, sun, and other elements of the solar system, and to build models of these relations. In another unit, students explore the motions of everyday objects and work with each other to develop ways of describing these motions. They wrestle with the difficult problem of how to describe the motions in a clear and consistent manner and ultimately raise deeper questions (explored in a later grade and unit) about how to explain these motions. In other units, students explore phenomena that involve them in theorizing about the nature of heat, matter, gravity, living things, heredity, and the origins of the universe.

These topics all involve areas where students' starting conceptions can be quite fuzzy and different from the ideas of science experts. Thus, in pursuing their investigations, students must work to clarify and understand their own initial ideas. They also encounter both anomalies that challenge their thinking and new ideas (from their teacher and others) that contribute to the process of conceptual change.

Most of the reform curricula showcased in the literature also stress the importance of picking problems that reveal important principles. However, these curricula vary in whether they focus on important design and engineering principles (Roth, 1996), on important methodologies for investigating a domain (Metz, 2000), or on underlying domain-specific theories that involve students in conceptual changes (Brown & Campione, 1994; Hennessey, in press; White, 1993). We believe that Hennessey's deliberate choice of problems that are on the frontiers of student understanding provides her students with a particularly rich opportunity to learn that ideas are multifaceted and involve explanatory conjectures that go beyond the information given. These problems also allow students to experience the difficulties in coming to understand their ideas and to learn about the kinds of mental work that go into understanding and clarifying ideas.

Representing ideas in multiple ways. A third feature of Hennessey's approach is her emphasis on having students take responsibility for representing their ideas in multiple ways. Clearly, if the focus of the curriculum is on the development and elaboration of student ideas, it is important to find ways to make those ideas public and open to inspection and debate. Not only do explicit representations help students concretize and systematize inherently abstract and complex ideas, they also help students clarify ideas or discover aspects of their ideas that are not clear to them. Finally, as has been well documented in the conceptual change literature, making ideas public facilitates the process of conceptual change itself (e.g., Hewson & Hewson, 1983; Minstrell, 1982; Smith et al., 1997).

Hennessey encourages students to use a variety of means to make their ideas public, including poster production, concept maps, physical models, drawings of conceptual models, word processing to write out ideas, audiotapes to dictate ideas, and small-group and whole-group discussion to present ideas orally. Poster production, audio recordings, and written statements not only serve to make ideas public, but also preserve a record of those ideas so that students can explicitly compare earlier and later ideas. Significantly, Hennessey encourages her elementary school students to represent, share, and analyze their ideas about domain-specific science concepts as well as their metaconceptual ideas about thinking, learning, and science. For example, students were asked to create concept maps of their notion of ideas and of the terms intelligible, plausible, and fruitful, and then to write word-processed essays in which they expressed their beliefs about the nature of learning and science.

Collegial learning communities and metacognitive discourse. The fourth and fifth features of Hennessey's teaching approach that we believe are central to enhancing the development of a constructivist epistemology among her students are the kinds of social and discourse structures that characterize her classroom. She has created what Brown and Campione (1994) called a *community of learners,* where social dialogue and collaboration is an essential aspect of the learning process, yet each student's voice is heard, respected, and valued. Students work together in a variety of ways—planning and conducting investigations; negotiating the meaning of words; learning to listen, share, and raise questions about each other's views—much in the way a community of scientists works together in developing and considering the viability of each other's ideas.

In such an environment, the teacher's role is complex: Often she serves as a facilitator and scaffolder of student inquiry. At other times, she introduces the views of members of the professional science community for her students to consider.[9] This kind of social environment facilitates students' awareness of the diversity of viewpoints and the ways in which they may (or may not) fully understand the ideas of self and others. It also widens the range of ideas students consider, which often leads them to develop more complex views. Given this kind of social environment, it is not surprising that students view these social interactions as vitally important to the learning and knowledge acquisition processes.

The collegial social environment in Hennessey's classroom calls for and is supported by an explicit metacognitive discourse among students about their ideas. In her own research, Hennessey (in press) extensively described the nature of this discourse and the variety of ways that the elementary students in her classroom are encouraged to develop metacognitive abilities. These include explicitly stating or identifying their own conceptions, considering the reasoning used to support a conception, considering the implications of a conception, temporarily bracketing or setting aside one's own conceptions to consider the competing views of others, reflecting on the status of conceptions of self and others (i.e., their intelligibility, plausibility, and fruitfulness), and evaluating the consistency and generalizability of a set of conceptions. She does not, of course, expect students to have all these metaconceptual skills initially. Rather, it is an explicit goal of her curriculum to help students build increasingly sophisticated metaconceptual skills and understandings over a 6-year period. (See Table 1 for a description of metaconceptual goals for students at each grade level and the way she ups the ante for students in Grades 4–6.)

---

[9]The views of science experts are presented as "one perspective among many" that are worthy of serious consideration and subject to rational debate; they are never dogmatically asserted as "the correct answer." They also are presented only after students have spent a fair amount of time exploring the conceptual terrain and developing their own views. In some cases, students consider competing views that exist (or have existed) within the science community.

In a recent observational study, Beeth and Hewson (1999a) described the complex kind of discourse that occurred in Hennessey's sixth-grade classroom during a 37-day unit on force and motion. They believed a crucial part of the artistry of her pedagogy is the way she and her students weave among three kinds of discourse throughout the unit: discourse about specific science concepts, metacognitive discourse concerning the status of their ideas, and discourse about epistemological standards. They also believed the depth of understanding her students achieve is significantly influenced by three sources of authority at play in this learning community: curricular authority, authority of epistemological standards, and personal authority. Curricular authority is in the hands of the teacher as she chooses concepts for study and the depth of target understandings. The authority of the scientific community's epistemological standards is introduced by the teacher when students are ready to apply them, by teacher-initiated questions such as these: Do you have evidence for your ideas? Are your ideas consistent with other ideas? Can you use your ideas to make predictions about new situations? Students also negotiate and apply their own epistemological standards. This personal authority is exercised by the students as they determine what to understand and ways to apply their ideas to new contexts. Hennessey respects and nurtures this personal authority by providing ample time for students to work with ideas, to negotiate standards for judging ideas, and to explore the status of ideas.

We believe that the discourse in Hennessey's classroom has all the elements of what van Zee and Minstrell (1997) called *reflective discourse*, a kind of discourse that they argued is crucial to supporting the process of conceptual change. In contrast to the teacher-controlled discourse of more traditional classrooms that follows the rapid-fire IRE format (teacher *initiates* question, student *responds*, teacher *evaluates* correctness of student response and then moves on to the next question and student), reflective discourse is more student centered, slower paced, and open ended. In particular, the questions and comments raised by the teacher or other students occur in reaction to student-initiated comments and often have the structure of a reflective toss (student utterance, teacher or student question or comment, student utterance). Such questions and comments may probe for clarification and elaboration of meanings, draw out a variety of views in a neutral manner, and encourage students to monitor the discussion and their own thinking. Both students and teacher take the important roles of questioners and commentators, and vigorous student–student–student reflective dialogues ensue.

Although reflective discourse with these features has been described for a variety of reform curricula both at the elementary and high school levels (Brown & Campione, 1994; Hennessey, in press; Lampert, 1990; Lehrer et al., 2000; Metz, 2000; Minstrell, 1982), there may be distinctive differences in the ways this type of discourse is orchestrated and the relative importance placed on its different components. For example, some teachers seem to focus on scaffolding discourse about evaluating an idea in relation to its fit with evidence (Brown & Campione, 1994;

Lehrer et al., 2000; Metz, 2000). Other teachers, such as Hennessey (in press), seem to put greater emphasis on scaffolding discourse about evaluating an idea in light of a variety of criteria: its intelligibility, its fit with their prior ideas, and its fit with evidence. The way these differences manifest in a given reflective discourse not only interacts with how the acts of inquiry occur in the classroom, but also may have implications for the epistemological lessons students learn from the curriculum in which the discourse evolves. Perhaps one reason that the students in Hennessey's classroom were so aware of the guiding role ideas play in scientific inquiry was the great emphasis she placed on having them evaluate their ideas not only in terms of fit with evidence, but also in terms of intelligibility and fit with their prior ideas.

Other features. A variety of additional factors may have contributed to the effectiveness of Hennessey's curriculum in bringing about change in student epistemological understandings. Hennessey is a knowledgeable scientist, with graduate study in the biological sciences. Her depth of scientific knowledge, as well as her willingness to research topics or contact experts, allows her to respond flexibly and intelligently to the questions and issues her students raise. She is knowledgeable about research on student conceptual frameworks and about reform efforts to teach science from a constructivist perspective, having completed doctoral work in science education. She, herself, has sophisticated epistemological views toward science. In addition, she is highly experienced at teaching elementary science from a constructivist perspective, having worked from this perspective with students over the last 20 years. Finally, by teaching science in a school that allows her to work with the same students over a 6-year period (and by having a student body that is relatively stable), she has an extended time scale, which makes it more likely that deep conceptual change can occur. She has unique opportunities to get to know her students and their thinking, to invite them to revisit and deepen their understanding of topics at varying points throughout the curriculum, and to remind them of their earlier views (e.g., she pulls out posters saved from prior years and discusses with students how their ideas have changed). In all of these respects, her classroom may represent a best case scenario for bringing about change in students' epistemological understandings.

However, we believe that we can learn a great deal from careful analysis of best case scenarios. They inform us of what is educationally possible given the prior concepts and developmental limits of elementary schoolchildren. What we learn is that elementary schoolchildren are much more capable of engaging with theory building and epistemological issues than many have assumed. Best case scenarios also can contribute to our understanding of exemplary educational practice and to our development of a more adequate vision of what the central goals of an elementary science curriculum can and should be. Like a number of other researchers in the field, we believe it may be particularly important to develop students' epistemological understandings early, as these views can provide an exciting and empowering framework to build on in their subsequent science educational experiences.

ACKNOWLEDGMENTS

This study was funded by McDonnell Foundation Grant 95–4 awarded to Carol Smith and Susan Carey.

We are grateful to Susan Carey for her help in planning this study and for her valuable comments on several drafts of the article. We are also grateful to the students we interviewed who, by so graciously and patiently answering our many questions, helped us to understand their perspectives on science. We thank the teachers and principals at the schools in which we interviewed, who were enthusiastic about this study and strongly supported our research efforts. Finally, we thank the editors and two anonymous reviewers whose perceptive comments, suggestions, and questions helped us to revise and greatly strengthen the article.

REFERENCES

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy (Project 2061)*. New York: Oxford University Press.

Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press.

Beeth, M. E. (1998). Teaching science in 5th grade: Instructional goals that support conceptual change. *Journal of Research in Science Teaching, 35*(10), 1091–1101.

Beeth, M. E., & Hewson, P. W. (1999a, March). *Facilitating learning of science content and scientific epistemology: Key elements in teaching for conceptual change*. Paper presented at the National Association for Research in Science Teaching, Boston.

Beeth, M. E., & Hewson, P. W. (1999b). Learning goals in an exemplary science teacher's practice: Cognitive and social factors in teaching for conceptual change. *Science Education, 83,* 738–760.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and educational practice* (pp. 229–270). Cambridge, MA: MIT/Bradford Press.

Carey, S. (1985a). Are children fundamentally different thinkers and learners from adults? In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills* (Vol. 2, pp. 485–517). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Carey, S. (1985b). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (1991). *The nature of science interview.* Unpublished manuscript.

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,* 235–251.

Chandler, M. (1987). The Othello effect: Essay on the emergence and eclipse of skeptical doubt. *Human Development, 30,* 137–159.

Chandler, M., Boyes, M., & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology, 50,* 370–395.

Chinn, C. A., & Malhotra, B. A. (in press). Epistemologically authentic scientific reasoning. In K. Crowley, C. D. Schunn, & T. Okada (Eds.), *Designing for science: Implications from professional, instructional, and everyday science.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30*(10), 1241–1257.

Clement, J., Brown, D., & Zietsman, A. (1989). Not all conceptions are misconceptions: Finding "anchoring conceptions" for grounding instruction on students' intuitions. *International Journal of Science Education, 11,* 554–565.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham, England: Open University Press.

Fabricius, W. V., Schwanenflugel, P. J., Kyllomen, P. C., Barclay, C. R., & Denton, S. M. (1989). Developing theories of the mind: Children's and adults' concepts of mental activities. *Child Development, 60,* 1278–1290.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchical skills. *Psychological Review, 87,* 477–531.

Gardner, H. (1999). *The disciplined mind*. New York: Simon & Schuster.

Gopnik, A., & Meltzoff, A. (1997). *Words and thoughts*. Cambridge, MA: MIT Press.

Grosslight, L., Unger, C., Jay, E., & Smith, C. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28,* 799–822.

Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction, 12,* 151–183.

Hennessey, M. G. (1994a). [Classroom essays by 6th-grade students about the nature of science or the nature of learning in Dr. Hennessey's science class]. Unpublished raw data set.

Hennessey, M. G. (1994b, May). *Conceptual change approach to learning science: The dynamic role of metacognition*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Anaheim, CA.

Hennessey, M. G. (1995, September). *Students' epistemological stance: Nature of learning and nature of science*. Presentation at the Cognitive Studies and Educational Practice Meetings of the McDonnell Foundation, Nashville, TN.

Hennessey, M. G. (in press). Probing the dimensions of metacognition: Implications for conceptual change teaching–learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hennessey, M. G., & Beeth, M. (1993, April). *Students' reflective thoughts about science content: A relationship to conceptual change learning.* Paper presented at the meeting of the American Educational Research Association, Atlanta, GA.

Herrenkohl, L. R., & Guerra, M. R. (1998). Participant structures, scientific discourse, and student engagement in fourth grade. *Cognition and Instruction, 16*(4), 433–475.

Hewson, M., & Hewson, P. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching, 20,* 731–743.

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

Hofer, B., & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88–140.

Honda, M. (1994). *Linguistic inquiry in the science classroom: "It is science, but it's not like a science problem in a book."* Cambridge, MA: MIT Working Papers in Linguistics.

Honda, M. (1996, February). *Developing an epistemology of science through linguistic inquiry*. Paper presented at the American Association for the Advancement of Science Annual Meeting and Science Innovation Exposition, Baltimore.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic.

King, P., & Kitchener, K. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.

Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills.* Orlando, FL: Academic.

Lampert, M. (1990). Connecting inventions with conventions. In L. Steffe & T. Wood (Eds.), *Transforming children's mathematics education* (pp. 253–265). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lederman, N. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29,* 331–359.

Lehrer, R., Carpenter, S., Schauble, L., & Putz, A. (2000). Designing classrooms that support inquiry. In J. Minstrell & E. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 80–99). Washington, DC: American Association for the Advancement of Science.

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

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

Minstrell, J. (1982). Explaining the 'at rest' condition of the object. *Physics Teacher, 20,* 10–14.

Montgomery, D. (1992). Young children's theory of knowing: The development of a folk epistemology. *Developmental Review, 12,* 410–430.

Nadeau, R., & Desautels, J. (1984). *Epistemology and the teaching of science*. Ottawa, Canada: Science Council of Canada.

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

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: Bradford Books/MIT Press.

Roth, W. -M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content and student responses. *Journal of Research in Science Teaching, 33,* 709–736.

Roth, W. -M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching, 30,* 127–152.

Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85,* 406–411.

Schwab, J. (1962). *The teaching of science as enquiry*. Cambridge, MA: Harvard University Press.

Smith, C., Maclin, D., Grosslight, L., & Davis, H. (1997). Teaching for understanding: A study of students' preinstruction theories of matter and a comparison of the effectiveness of two approaches to teaching about matter and density. *Cognition and Instruction, 15,* 317–394.

Sodian, B., & Schrempp, I. (1997, March). *Metaconceptual knowledge and the development of scientific reasoning skills*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development, 58,* 424–33.

Songer, N. B., & Linn, M. C. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28*(9), 761–784.

Strike, K., & Posner, G. (1985). A conceptual change view of learning and understanding. In L. West & A. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 211–231). New York: Academic.

Taylor, M. (1988). Conceptual perspective-taking: Children's abilitiy to distinguish what they know from what they see. *Child Development, 59,* 703–718.

van Zee, E., & Minstrell, J. (1997). Using questioning to guide student thinking. *The Journal of the Learning Sciences, 6,* 227–269.
Wellman, H. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.
White, B. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10*, 1–100.
Wiske, M. S. (1997). *Teaching for understanding*. San Francisco: Jossey-Bass.

APPENDIX A
NATURE OF SCIENCE INTERVIEW

Goals of Science

1. What do you think science is all about?
2. What do you think the goal of science is?
3. What do you think scientists do?
3a. How do they achieve the goals of science?

Types of Questions

4. Do you think scientists ask questions? What sorts of questions? IF NO, go to Question 6.
5. How do scientists answer their questions?
5a. Can you give an example of a scientist's question and what he or she would do to answer it?

Nature and Purpose of Experiments

6. What is an experiment?
7. Do scientists do experiments? IF NO, skip to Question 10.
8. Why do scientists do experiments? IF "to test ideas," THEN: How does the test tell the scientist something about the idea?

Role of Ideas: Conceptions of Hypotheses and Theories

9. How does a scientist decide what experiment to do?
10. Have you ever heard the word *hypothesis*? IF NO, explain: A hypothesis is an idea scientists have, an idea about how an experiment would turn out. IF YES, ask: What is a hypothesis? IF "educated guess" or "guess" THEN ask: Do you think a hypothesis is the same as a guess or do you think that there is a difference? What is the difference?
11. Do you think a scientist's ideas influence the experiments he or she does? IF YES: How? IF NO: Do scientists ever test their ideas?

12. How do you think scientists come up with their ideas?
13. Have you ever heard the word *theory*? IF YES: What is a theory? Do you think scientists have theories? IN ALL CASES, EXPLAIN: "A theory is a general idea about how and why things happen the way they do. For example, biology is a theory about living things."
14. Do you think a scientist's theory influences his or her ideas about specific experiments? How?

Unexpected Results and Disproving Ideas

15. If a scientist does an experiment and the results are not as he or she expected, would the scientist consider this a bad result? Why or why not? Can they learn anything from this? What?
16. Say a scientist is going to do an experiment to test his or her idea. Would a scientist do an experiment that might prove this idea is wrong? Why?

Nature of Change Processes

17. What happens to the scientists' ideas once they have done a test?
18. Do scientists ever change their ideas? IF YES: When would they do that and why?
19. Do scientists ever change their whole theories? IF YES: When and why?

Achieving Goals and Making Mistakes

20. Do scientists always achieve their goals? If not, why not?
21. Can scientists make mistakes or be wrong? How?

APPENDIX B
CODING GUIDE FOR THE FOUR QUESTION CLUSTERS

Cluster 1: Goals of Science

Level 1

The core feature of Level 1 answers about the goals of science is that students talk of scientists' ideas, experiments, and experimental results in an undifferentiated fashion.

Doing things. At Level 1a the goals are simply the concrete activities and products of science. There is no awareness of the role of scientists' ideas in guiding those activities or devising those products. For example, scientists "work in labs," "do experiments," "try things out to see if they work," "invent things," or "find cures."

Gathering information. At Level 1b, the goals involve finding or discovering new information. The information is conceptualized as out there to be discovered. There is no differentiation between scientists' ideas and their observations or experimental results, and there is no awareness of the role of scientists' ideas in guiding the search process. For example, students talk of "making new discoveries," "learning new things," "finding answers," or "solving problems" with no mention of scientists' ideas playing any role in these processes.

Level 1.5

As students make the transition from Level 1 to Level 2, they become increasingly aware of the role of ideas and thinking about ideas in science, although the nature of the ideas and kind of thinking still remains vague and ambiguous.

Thinking about ideas or data. These students begin to be aware that scientists have ideas that affect their work. For example, scientists "think about their ideas" or "think about what their data means." They are vague, however, about the nature of scientists' ideas and do not go so far as to see that these ideas are being tested in experiments.

Finding how it works, unelaborated. These students mention that scientists are concerned with finding out "how something works" but give no further example of what they mean. Hence it is ambiguous whether they are referring to a concrete procedure (i.e., how to work something) or an underlying mechanism.

Level 2

The core feature of Level 2 is that students have made a differentiation between scientists' ideas, experiments, and experimental results. This differentiation enables students to have a notion of explanation (i.e., ideas that explain some phenomena or test results) and hypothesis testing (i.e., evaluating an initial idea in light of results). As they begin to realize that ideas are complex and have parts, they also realize that ideas take work to understand and develop over time. Students may comment on any one of the following four ideas.

Finding explanations. These students focus on scientists' concern with finding or figuring out (a) how something works (with a clear example showing that they mean some underlying mechanism) or (b) why it happens.

Testing ideas. These students make clear that the scientist has some initial idea that is being tested by the experiment (e.g., he "tests his ideas by doing experiments"; he "does experiments to find out if his ideas are right or wrong").

Understanding ideas. These students see the goal of science as understanding one's ideas, where it is assumed that ideas are not immediately graspable, and that it may take some work to understand them. For example, scientists "try to understand their ideas and the ideas of others," "they try to figure out what their ideas mean."

Developing ideas. These students see the goal of science as developing one's ideas, where it is assumed that scientists build, alter, add to, or take away from ideas.

Level 2.5

At Level 2.5, some students show a more sophisticated understanding of the goals of science by interrelating two or more Level 2 ideas in their response. For example, the goal of science is to explain how things work and to test those explanations. Other students articulate a more sophisticated idea about how the processes of testing and development occur. More sophisticated ideas include the following.

Testing ideas: Need for "fit." These students talk of assessing the "fit" between one's initially held ideas and a pattern of evidence. In so doing, they show a more complex understanding of the testing process (i.e., it involves putting together multiple pieces of evidence, not just finding out if an initial idea is right or wrong based on a one-shot experiment).

Developing ideas: Revision or dissatisfaction. These students give a more elaborated discussion of the process of developing ideas than those who are scored at Level 2. For example, they might talk of development as involving the revision of an idea or as involving a process of sensing dissatisfaction with an idea. These comments show they are beginning to understand the development of ideas as a complex, multistepped process, not just a simple process of adding to existing ideas.

All of these ideas are considered Level 2.5 rather than Level 3 because students do not discuss the development of increasingly more explanatory theories as the main goal of science, a process that brings together and integrates the three Level 2 notions of development, testing, and explanation.

Cluster 2: Types of Questions

Level 1

Again the core feature of Level 1 is that students do not clearly differentiate between scientists' ideas and experiments. Hence, scientists ask the following types of questions.

Procedural questions. These questions concern how to do something (e.g., how to cure something, make a machine, carry out an experiment).

Journalistic questions. These questions concern either concrete observables (i.e., what happened, who did it, when, where) or people's opinion about what happened.

Level 1.5

At this transitional level, students mention questions that are beginning to be more sophisticated in one of several ways, but still treat them fairly concretely.

Variable relation questions. These questions are classic science fair questions such as, "Does listening to music affect the time it takes to do your homework? Does one kind of gas give you better mileage?" Although these questions involve comparing the performance of two groups, they still focus on easily observable variables with no attention to the notion of an underlying mechanism.

Unelaborated how it works questions. These students mention that scientists are concerned with finding or figuring out "how something works" but give no further example of what they mean.

Questions about theoretical entities (treated factually). These students ask potentially deeper scientific questions (e.g., questions about the shape of the earth or the shape of atoms), but then go on to imply that these questions could be answered in a simple fashion via direct observation.

Metacognitive questions about Level 1 issues. These students imagine that scientists would pose metacognitive questions about their goals and knowledge states (e.g., Why am I doing this experiment? What do I know and what don't I know?). They go beyond being simple Level 1 questions because they are reflective in nature, but they still concern primarily Level 1 issues (e.g., what one is doing, what one knows).

Level 2

At Level 2, students are more aware of deeper scientific questions: questions about how the world works; questions involving theoretical terms, not just observables. They also ask more sophisticated reflective questions.

Explanation questions. These are basic questions about how something works (accompanied by a clear example) or why something happens. Included in this category are "why" questions in which students may be looking for either a teleological or a mechanistic explanation (e.g., Why did God create trees? What causes disease?).

Questions about theoretical entities (treated more abstractly). These questions are about some unseen entity or abstract concept (e.g., questions about the shape of the earth, the composition of the atom, the nature of gravity). Furthermore, students either explicitly acknowledge that they are hard to answer, or, at least, do not foreclose this possibility in their responses.

Metacognitive questions about Level 2 issues. These metacognitive questions probe the reasons for one's ideas and for the quality of one's ideas (e.g., Why do I think that? or Is my idea intelligible or plausible?).

Level 2.5

At Level 2.5, students show greater awareness that the type of questions scientists ask are complex and not simple to answer by interrelating two or more Level 2 ideas. For example, they say that scientists ask questions about how things work or why things happen, and then go on to give a specific example that makes reference to an unseen theoretical entity. In so doing, they show an appreciation that explanations in science appeal to deeper theoretical terms. Or, they combine mention of either an explanatory or a theoretical question with a sophisticated metacognitive question about the intelligibility or plausibility of their ideas. In this way, they show an understanding that the ideas being investigated are complex and take work to understand.

These answers fall short of Level 3 understanding in that they do not discuss how multiple levels of questions in science interact and support each other. For example, scientists may ask a general theoretical question, but they need to give more specific operational definitions (and answer questions about the relations among measurable variables) as part of answering that broader question.

Cluster 3: Nature and Purpose of Experiments

Level 1

At Level 1, students think of experiments in practical terms—as producing a desirable outcome or a new fact. This notion reflects an underlying conflation of ideas and either experiments or experimental results.

Try out or find cures. The motive for science experiments is essentially practical: to try something to see if it works or to find cures.

Find answers. The motive for science experiments is to find answers, solve problems, or to learn something new.

Level 1.5

At this transitional stage, students mention that experiments are done to find out how something happens, but they do not elaborate on what they mean by how something happens.

Level 2

At Level 2, students think of experiments as tests of ideas or as more generally contributing to search for explanations or the development of ideas. Examples include the following.

Find explanations. The motive for science experiments is to find an idea that explains how something works (with elaboration) or what causes something to happen. Note that students talk about "finding" rather than "testing" explanations.

Test ideas. The motive for doing experiments is to test initially held ideas. Note that the nature of the idea being tested is generally left ambiguous, but there is evidence that the student thinks the scientist had the idea prior to doing the experiment.

Develop ideas. An experiment is a method of developing initially held ideas, or of making an idea better, clearer, or more understandable to self and others.

Level 2.5

At Level 2.5, students interrelate two or more Level 2 ideas, and thus show a more sophisticated understanding of the purpose of experiments. For example, they talk of experiments not only as a means of testing ideas but also as part of a larger process of developing or improving ideas. They could also describe experiments as a way of testing causal ideas.

These responses fall short of Level 3 understandings in that students do not discuss why experiments enable causal ideas to be tested (i.e., because experiments are designed to generate data that empirically distinguish alternative causal hypotheses and involve the controlled manipulation and comparison of variables), nor do they discuss hypothesis testing as a means of providing indirect evidence for or against a larger theory.

Cluster 4: Nature of Change Processes

Level 1

At Level 1, students conflate ideas and experiments or talk of keeping or abandoning an idea based on a simple observation or a single experiment. The ideas they discuss are, by implication, simple ideas because they mention no process of development or elaboration.

Level 1.5

At Level 1.5, students are beginning to be aware that thought or effort go into changing ideas. For example, students mention that scientists might change an idea after thinking about what went wrong or why they did not get the desired result, or after thinking about what might be a better idea. They might also mention that scientists need to repeat the experiment (to make sure they are really wrong) before changing their ideas, or they might mention that scientists would be reluctant to change their whole idea after a single experiment.

Level 2

At Level 2, students are aware that the change process either (a) involves the development of ideas (not just the abandonment of ideas) or (b) involves finding better explanations, using complex evidence, or being constrained by one's prior ideas.

Development of ideas. These students talk of change as involving the development of ideas, rather than the keeping or abandoning of a (static) idea based on the results of an experiment. They generally think of ideas as having parts. Development involves either (a) simply adding to (or elaborating) ideas over time, or (b) keeping some parts and changing others. In either case, the process of change involves some sustained effort over time.

Complex evidence. Students who are scored as understanding that change involves complex evidence discuss the need for multiple experiments even in the face of positive results (i.e., the need to "go at" the same idea in several ways to confirm or disconfirm it, or the need to have a fit between the idea and a pattern of results).

Better explanation. Students show an awareness that change involves finding a better explanation in one of two ways: (a) by making a metastatement that one changes an idea when one finds a better explanation (without an explicit example of what they mean by a change in explanation), or (b) by giving an example that involves changing some deep idea or explanation (without explicitly commenting that it is an example of changing an explanation).

Constrained by prior ideas. Students can give evidence that they think the process of changing an idea is constrained by prior ideas in two ways: (a) by talking about the "need for new ideas to make sense," or (b) by making a metastatement that it is hard to abandon a belief if it is something that one strongly believes in.

Level 2.5

At Level 2.5, students show a more sophisticated understanding of the complexity of change in that they realize that change is constrained both by patterns of evidence and by one's ideas or general search for a "best explanation." That is, they combine complex evidence themes with better explanation or constrained by ideas themes. Alternatively, they articulate one of the following more sophisticated ideas about how change involves finding a "better explanation" or is "constrained by prior ideas."

Better explanation: Gives an example and a metastatement. These students combine metaconceptual talk of finding a better explanation with a clear example of deep theory change.

Constrained by prior ideas: (a) Need for fit and coherence among ideas or (b) prior ideas affect interpretations. In (a), students go beyond simply mentioning that new ideas need to "make sense" by explicitly talking about how ideas need to fit together and be coherent and consistent with each other. In (b), students go beyond asserting that strongly held beliefs can be difficult to change by explaining the reason for the difficulty: These beliefs affect the very process of observing and interpreting results.

These Level 2.5 responses fall short of Level 3 responses because students do not discuss how theories are frameworks that constrain the hypothesis generation process nor do they distinguish between the change processes of normal science (in which new hypotheses are developed that elaborate a given theoretical framework) and the change processes of revolutionary science (in which the general theoretical framework is challenged).

APPENDIX C
CODING GUIDE FOR IDEAS ABOUT SOCIAL INTERACTION

Level 1

At Level 1, some students see scientists as engaging in concrete activities together (e.g., they work together, improve each other's technology, help each other choose materials and procedures, run experiments together). Other students view knowledge as certain (as opposed to conjectural) so they depict scientists as exchanging factual information with each other (e.g., they share facts, knowledge, answers, or results, and ask each other informational questions).

Level 1.5

At Level 1.5 there is a notion that a range of divergent viewpoints exist, but ideas are not considered conjectural. Ideas are equally valid beliefs differentiated only on the level of factual information. Students see scientists as exchanging and comparing ideas to enlarge the scope of ideas they consider, but not to evaluate, develop, or make sense of ideas. For example, scientists share thoughts, do experiments to answer each other's questions, compare ideas about results, and disagree about what is right without any goal of or process for achieving agreement.

Level 2

At Level 2, students view scientists as engaging in processes such as explaining phenomena, or evaluating, understanding, clarifying, and developing ideas, processes that imply a grasp of ideas as conjectural. For example, scientists develop ideas together by using parts of different people's ideas to build a new idea. They exchange explanatory ideas about how things work. They strive to understand each other's ideas[10] (or, in more specific cases, they judge the status of each other's ideas

[10]Note that the term *each other* is used in this coding guide to refer both to explicit comments on how social interaction enhances both parties' ideas and to comments on how social interaction enhances the ideas of one of the parties.

as intelligible, plausible, useful, etc.) or they use each other's ideas to clarify an idea. They test each other's ideas with Experiments in an effort to prove them right to each other. Finally, they test each other's ideas using a social process; that is, they assess the viability of an idea by measuring the level of agreement among a group of colleagues who are engaged in a common investigation (note that this is a distinct alternative to carrying out evaluation with respect to evidence).

Level 2.5

Students at Level 2.5 show a deepening understanding that ideas are conjectural, interpretive, and uncertain in nature by depicting scientists putting substantial time, care, and effort into the challenges of persuading others, integrating diverse ideas, or making decisions about the viability of ideas and admissibility of evidence.

A Level 2.5 composite response interrelates two Level 2 ideas in a social context (e.g., scientists develop ideas using each other's plausible ideas, or they develop an explanatory idea by putting together ideas from several scientists). The following are more complex Level 2.5 ideas.

Using evidence to establish the viability of an idea to others depicts scientists as using evidence to convince others that an idea is either wrong or plausible, or to support an idea so well that it is resistant to being proven wrong by others.

Operating with social principles depicts scientists as upholding the responsibility they have to each other and to the scientific community to admit only ideas for which one has evidence into the social arena for review.

Building consensus about ideas focuses on the critical role of cooperation in consensus building; it emphasizes the need for colleagues to work together as a group to formulate, confirm, and disconfirm ideas, or to evaluate and develop an idea in terms of fit with each other's well-confirmed ideas.

Influencing each other's interpretations emphasizes the role scientists take in helping each other to minimize personal bias (e.g., if a prior idea affects how one scientist sees a phenomenon, it may take another scientist to convince him that his interpretation is mistaken).

These responses fall short of the Level 3 understanding that ideas are embedded in theoretical frameworks and that this embeddedness makes it difficult to achieve consensus about what is the most viable explanation.