Chapter 4
What Can Argumentation Tell Us About Epistemology?

William A. Sandoval and Kelli A. Millwood

Who, besides scientists, engages in what we would call scientific argumentation? When? For what purpose? As calls for argumentation to take a central place in science instruction increase (Driver et al., 2000; Duschl & Osborne, 2002; Kuhn, 1993b), answers to these questions become more important. There are two key claims for engaging students in scientific argumentation. One is that argumentation is a central practice of science, and thus should be at the core of science education. The other is that understanding the norms of scientific argumentation can lead students to understand the epistemological bases of scientific practice. We are more interested in this second claim. We think it unlikely that people who do not practice science are likely to engage in truly scientific argumentation. At the same time, we see everyday contexts all around us where people might apply scientific arguments to further other kinds of arguments. For example, using arguments about global climate change to argue for or against particular energy policies or even personal consumer decisions.

Consequently, the focus of our studies has been to understand how students' practices of scientific argumentation reflect their understanding about science: about what makes a claim scientific, and how such criteria are related to methods that scientists use to generate and to warrant claims. Thus, our studies of students' efforts at scientific argumentation are aimed at helping us to understand students' epistemological ideas about science. Hence, what can argumentation tell us about epistemology?

Epistemology and Practice

Epistemology is the branch of philosophy concerned with the study of knowledge. Philosophers of science have been concerned with outlining an epistemology of science—the logical and philosophical grounds upon which scientific claims are advanced and justified. This move itself presupposes that scientific knowledge and the processes of its construction are potentially different from other forms of knowledge and knowing. Scientific epistemology is a description of the nature of scientific knowledge, including the sources of such knowledge, its truth value,

S. Erduran and M. P. Jiménez-Aleixandre (eds.), *Argumentation in Science Education.*

scientifically appropriate warrants, and so forth. Psychologists take this notion of epistemology and internalize it, defining personal epistemology as the set of beliefs that individuals hold about the nature of knowledge and its production. Thus, psychologists speak of the scientific epistemologies held by individuals. In science education, research into students' epistemological ideas has occurred under the name of NOS (Nature of Science) research. The move to studying the epistemological ideas that students may have about science by studying how they make scientific arguments is quite recent.

Epistemics of Argumentation

One of the aims of research on argumentation in science education is to get students to argue like scientists. Broadly, the goal is to get students to use evidence to support claims that they make. Clearly, this coordination of claims and evidence raises inherently epistemological questions. What counts as a claim? What counts as evidence? How do you decide what sort of evidence supports, or refutes, a particular claim? How are individual claims organized to produce a coherent argument? What kinds of coordination of claims and evidence make an argument persuasive? How one answers these questions through a specific argument, whether consciously asked or not, may reflect epistemological notions about claims, evidence, and other forms of knowledge and their production.

Student argumentation has been studied across a range of age levels in two basic contexts, oral and written argumentation (see Duschl, this book; Kelly et al., this book). Oral argumentation has almost exclusively been studied within contexts of collaborative inquiry or problem-solving. Researchers have construed the dialogue that students engage in during such collaborations as argument, and analyses have thus focused on the epistemic moves that students make during such conversations. One finding that has emerged from these studies is that students commonly advance claims without providing explicit justifications (or warrants) for those claims (Erduran et al., 2004; Jiménez-Aleixandre et al., 2000; Kelly et al., 1998; Resnick et al., 1993). These studies show that claims are justified only when they are challenged, and even then not always. Claims are often offered without relation to other elements in an ongoing argument (Jiménez-Aleixandre et al., 2000). Furthermore, students can provide warrants for claims in a number of ways, including appeals to both empirical evidence and hypothetical or theoretical ideas (Kelly et al., 1998).

Argumentation has also been studied through examinations of student writing, and again across a number of age and grade levels. As with analyses of oral argumentation, analyses tend to focus almost exclusively on the structure of students arguments, with various efforts to try and capture argument quality. As with studies of oral argumentation (e.g., Erduran et al., 2004; Jiménez-Aleixandre et al., 2000), researchers have applied Toulmin's (1958) argument structure (e.g., Bell & Linn, 2000). Kelly and Takao have developed a scheme of assigning "epistemic levels" to students' written arguments that distinguish increasing levels of epistemic complexity in student writing (Kelly & Takao, 2002; Takao & Kelly, 2003). In our own prior work, we have examined the levels of empirical evidence that students provide to justify particular claims (Sandoval, 2003; Sandoval & Millwood, 2005).

Most of these studies show how students fail to produce sufficiently scientific arguments through their writing, while at the same time, writing arguments seems to help students learn important scientific ideas (Bell & Linn, 2000; Sandoval, 2003). Students seem to have similar issues in writing arguments as they do in oral argument. For instance, students often fail to provide sufficient warrants for their written claims (Sandoval, 2003; Sandoval & Millwood, 2005). Further, they often fail to make explicit the links between data and claims about data, a finding common to Sandoval's studies and to those by Kelly and Takao (Kelly & Takao, 2002; Takao & Kelly, 2003).

Taken as a whole, research on students' practices of argumentation suggests the complexity of appropriately coordinating causal claims with evidence. Of course, at issue is what exactly it means to coordinate claims and evidence appropriately. For instance, within professional science contexts the claims that need to be explicitly warranted are only those that have yet to be accepted. Thus, students' failure to warrant particular claims is not inherently "unscientific," but may simply reflect their belief that those claims are already believed. The analyses of oral argumentation mentioned above show that students do, in fact, provide warrants when claims are contested, as scientists themselves do. Issues of claim-evidence coordination in written arguments, on the other hand, may indicate that students do not see the rhetorical task of a scientific argument as one of persuasion. The tight coupling of evidence and claims that we take for granted in scientific arguments reflects a rhetorical effort to persuade readers of the preferability of an argument. Such rhetorical strategies are not simply social, but may be necessary to make novel ideas comprehensible (Kitcher, 1991).

Limits to "Practice Studies"

This is to say that the kinds of argumentation that students do, and do not, perform in school science contexts are some messy reflection of epistemological ideas they may have about the nature of claims (about the natural world) and the kinds of evidence or other justifications that make such claims believable. We say messy, however, because there are a number of other factors, besides epistemological, that might influence student argumentation. Consequently, while studies of students' practices of argumentation are important, they are not sufficient to help us understand the epistemological ideas that students bring to bear during such work, or how such work develops those ideas. Even studies that look at students' practices in detail have to make quite speculative inferences on how students interpret the purposes of their activity (cf., Sandoval & Morrison, 2003).

Kelly in particular has argued that the focus of research should be on students' sense making practices, either through traditional or inquiry-oriented instruction (Kelly et al., 1998; Kelly & Duschl, 2002). Certainly, the development of certain practices that can be labeled as scientific is a main goal of recent reforms. Kelly's perspective on science is strongly influenced by science and technology studies (STS), which argue that practice is the key feature to emphasize in science because science is a practice. Yet, most students will not really engage in science as a practice. Rather, as citizens they must be able to reflect upon scientific knowledge claims as they relate to personal or policy decisions. It is far from clear that simply engaging in practices of authentic science leads to such reflective ability. In fact, available evidence suggests that this is unlikely (see Sandoval, 2005). Studies of practice in themselves do not provide enough of a window onto students' epistemological ideas about science, because there are many possible ideas that might motivate particular practices.

We have recently laid out a theoretical perspective on epistemological beliefs that we call practical epistemologies (Sandoval, 2005). This theory proposes that students develop highly contextualized epistemological ideas as a result of their practical experiences trying to understand and explain the world they live in. We further propose that such epistemological beliefs drive, at least in part, people's efforts to explain new situations. That is, one's ideas about what counts as a satisfactory explanation—plausibility, fit with what you already know, standards of evidence, etc.—influence one's attempts to create explanations. This theory is an attempt to explain the evidence (reviewed by Sandoval, 2005) that students' epistemological beliefs appear to vary, in content and apparent sophistication, depending upon the context in which they are elicited, as well as the apparent paradox that students' practices of science are often more sophisticated than their expressed beliefs about science. Our view is similar to other proposals that epistemological beliefs are fragmentary and deployed as resources (Hammer & Elby, 2002) or repertoires (Bell & Linn, 2002). We believe that one promising way to investigate how such practical epistemological beliefs influence argumentation, and science learning more generally, is to augment studies of students' practice with techniques to identify epistemological beliefs directly.

Students' Ideas about Warranting Claims

Here we describe a recent study we undertook to test the viability of the practical epistemologies theory. We decided to look specifically at students' ideas about how to warrant claims. We started with such a project because of the historical interest both in science education and developmental psychology in how students coordinate claims with evidence (see for instance Kelly et al., this book; Garcia-Mila & Andersen, this book). We also expected that whatever ideas students hold about how they know particular claims, and why they believe them, could be straightforwardly assessed by asking students about them. That is, while asking students about professional science has often been problematic (see Sandoval, 2005), we expected that students could explain their own work.

Setting

We conducted this study in a Grade 7 classroom in an urban middle school in Los Angeles; 33 students participated (20 boys, 13 girls), with their teacher. This school is in a middle-income community of the city, with a study population that is 75% Caucasian, 14% Latino, 10% Asian-American, and 1% African-American; and 12% of the students received free or reduced lunches.

We explored students' ideas about warranting claims within the context of a three week science unit called *Sensing the Environment* (Griffis & Wise, 2005). This unit explores plant adaptation to the environment, focusing on topics of photosynthesis and transpiration and plants' evolutionary adaptations to climates to manage these processes. This unit is part of a curriculum development project that is producing curriculum units that frame student inquiry around data sets collected by remote sensor networks developed and deployed by the Center for Embedded Networked Sensing (CENS, cens.ucla.edu). One of the aims of the project is to study how students' argumentation practices develop through scaffolded instructional experiences, and to examine the effects that such inquiry may have on their ideas about the nature of science.

Students began the unit by looking at a picture of some local mountains and being asked what they noticed (a version of this picture is shown on the right of Fig. 4.1). Students noticed that the plants in the photo look different from each other, and in this way the teacher posed the question, "Why do plants look different?" This provided the guiding question for the unit. Following a series of laboratory activities in which students explored photosynthesis and transpiration, students finished the unit by conducting an investigation into the relations between plant leaf structure (mainly size) and environmental factors using an online environment shown in Fig. 4.1. Students worked in groups of two or three to explore how differences in temperature, humidity, and light intensity affected leaf size.

The online investigation involved a series of steps through which students made decisions about what data to analyze. From a series of pull down menus, students decided which of the three environmentally different areas to investigate, how to aggregate the data (average, maximum, minimum), and over what time frame (hourly, daily, or weekly). They were then able to generate graphs of the type shown in Fig. 4.1. To the right of the graph is a leaf gallery depicting images of leaves in area 1. Beneath the graph is a comparison option that allowed students to add an additional area (station) or environmental variable to the graph.

In addition to comparing environmental variables in different areas graphically, students could calculate the average leaf size within an area. By clicking on the individual leaves on the screen (e.g., in area 1, Fig. 4.1), students could obtain an enlarged image of the selected leaf in addition to its name and surface area. In order

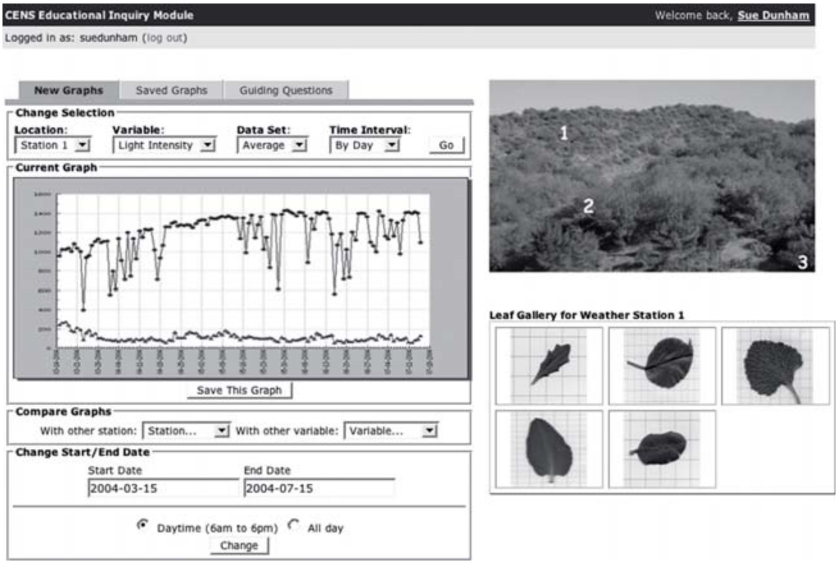

Fig. 4.1 Data query interface to Sensing the Environment online tool

to find the average surface area of all the leaves in a particular area, students had to manually average the various surface areas of the leaves. They could then use this data in addition to the graphs they created to look for patterns between leaf size and environmental trends.

Students looked for patterns in these data to figure out why plants look different. An example of a pattern that a group of students noticed is that the leaves in area 1 are smaller on average than the leaves in area 3, and the temperature in area 1 is higher than the temperature in area 3. These data can be used to support the claim that plants have smaller leaves in hotter areas.

A separate tab in the interface presented students with a series of questions to guide their investigations. These questions helped students to generate causal ideas about why plants look different and also aided them in writing their final explanations. Students were instructed to answer these questions as they proceeded through their investigations. After their collaborative investigations, which lasted three classroom periods for approximately 135 minutes, individual students synthesized the data they had analyzed to write an explanation for the question of why plants look different. Individual students were able to print out their group responses to the guiding questions and then write an explanation using the information from the guiding questions. Before students wrote their explanation, they were given a rubric and they discussed what elements should be included in the explanation. After writing their essays, students reviewed their own and three of their peers' explanations, using this rubric.

The rubric students were given, designed by study teachers, provided a highly structured set of rules for their essays. These included repeating the guiding questions that students had been given in the online environment, with instructions to make sure that they had been answered. These were followed with instructions for what is known in the United States as the common five-paragraph essay: an introductory paragraph, three body paragraphs that should address each of the climate areas investigated, and a concluding paragraph where students were instructed to make a general claim about why plants look different. Additionally, students were instructed that their essays should demonstrate that they: had learned key concepts; identified appropriate variables; searched the data for “patterns, anomalies, relationships”; and presented data-based arguments. We point out that in this version of the unit there was no discussion of just what, exactly, a “data-based argument” might look like. In this sense, this study is a replication of an earlier study (Sandoval & Millwood, 2005), and provides a baseline understanding of students’ evidentiary standards.

What Students Do and What They Say

We used a variety of methods to understand both students’ practices of warranting claims and their ideas about that practice. We analyzed the contents of students written essays to determine how they warranted their written claims. Following the end of the unit, each of the 33 students in the class was individually interviewed about their essay. Finally, students completed a version of the Views of Nature of Science (VNOS) questionnaire (Lederman et al., 2002) modified for students of this age range, known as POSE (Perspectives On Scientific Epistemology; Abd-El-Khalick, 2002). Our aim in triangulating across these data sources was to answer two questions. First, did students express beliefs about their own written work that might explain their actual practices of using data to warrant claims? Second, were these expressed beliefs about their own work related to their expressed ideas about how scientists warrant claims?

Warrants in Written Arguments

We collected the essays written by all 33 students in this classroom. Students’ explanations were analyzed with a scheme adapted from Sandoval and Millwood (2005). The first part determined if the student articulated the four main conceptual claims targeted through the unit: the function of a leaf, the structure of a leaf, environmental variations that affect leaf structure and function, and the differential fitness of structural (size) variations in leaves under varied environmental conditions. A student could score one point for each articulated claim, for a maximum of four points. As with our earlier analyses, students could receive a point for articulating a claim as long as the stated claim could be interpreted as pertaining to one of the four ideas in the coding scheme. Such claims did not have to be correct, however. Our aim was not to see how many students got the right answer, but to understand how students warranted the claims that they made.

The second part of the scheme analyzed the warrants students provided for their articulated claims. All of the claims, except function, had a four-level scheme to assess how well the student warranted their claim. Warrants for claims about leaf function were not scored for two reasons. One, leaf functions—photosynthesis and transpiration—were extensively explored prior to the online investigation and could safely be taken as shared knowledge in the classroom and consequently not likely to need a warrant (cf., Latour, 1987, and his analysis of warrants in scientific articles). Second, there were no available data in the online environment to support or refute particular claims for leaf function, so students had no opportunity to provide specific warrants for such claims. For the other claims in our argument structure, students could receive a score from 0—providing no warrant, to 3—providing a "full" warrant, as described in Table 4.1. A higher score indicated that students' warrants for claims were more like scientists'. Note that claims for differential fitness were not warranted through citing data, but by citing the appropriate scientific principle that explains the available data—what is known as the photosynthesis–transpiration compromise: the more surface area a leaf has, the more photosynthesis it can do. At the same time, however, a larger surface area increases the rate of transpiration (the process by which plants draw water through their roots—as water evaporates through the leaf surface this creates a vacuum pump to draw water through the plant). If a plant transpires too quickly in a hot or sunny climate it can dry out and burn. Since students had explored the processes of photosynthesis and transpiration independently prior to this investigation, we expected that at least some of them should be able to use the compromise as a justification for their claims.

While we did not require that the claims that students articulated be correct, our scoring of warrants required that warrants be appropriate for each claim. For example, if a student made a claim about leaves in area 2 being the biggest, but only

Table 4.1 Coding Scheme Used to Score Warrants in Students' Written Arguments

	Function	Structure	Environmental Variation	Differential Fitness
0	N/A	No relevant warrant given		
1		Leaf size data given for only one area	one environmental variable given, from only one area	Use principle of either transpiration or photosynthesis
2		Leaf size given for two areas	two or more variables compared from two or more areas	Use principle of photosynthesis–transpiration compromise
3		Leaf size given for all three areas	All three variables systematically compared across all three areas	Use photosynthesis–transpiration compromise and principle that plant will try to maximize leaf size

provided data about the leaf sizes in area 1 that would be scored as 0 for the structure claim. Similarly, if a student made a claim about the leaves in one area being larger than those in another, but the data showed that, in fact, those leaves were smaller, then that claim would be scored as unwarranted (a score of 0).

As already mentioned, we collected and scored all of the individual essays written by students in this class, 33 in all. The combined argument score could range from 0 to 13 (4 points for articulating all target claims, plus 9 points for fully warranting the last 3 claims). The mean total score for explanations was 6.30 (SD = 3.40). The mean articulation score was 2.78 (SD = 1.22) and the mean warrant score was 1.95 (SD = 2.45). The majority of students articulated all of the claims, but most students did not provide warrants for their claims (see Fig. 4.2). Approximately 79% of the students provided a warrant for their structure claim, and most of these were full warrants—comparing leaf sizes across all three available areas. Yet, only 33% of the students provided a warrant for environmental variation, and 18% of the students provided a warrant for differential fitness.

In contrast to our previous study (Sandoval & Millwood, 2005), students here were not likely to warrant the claims they made in their explanations, despite a similar level of instruction about the need for data to support claims. It may be that there is some developmental explanation, as the students in this study were slightly younger (by 2 years, their age being 12–13) than the students in that previous study. Given the evidence available of much younger children using evidence to support claims (Lehrer & Schauble, 2006), this explanation seems unlikely.

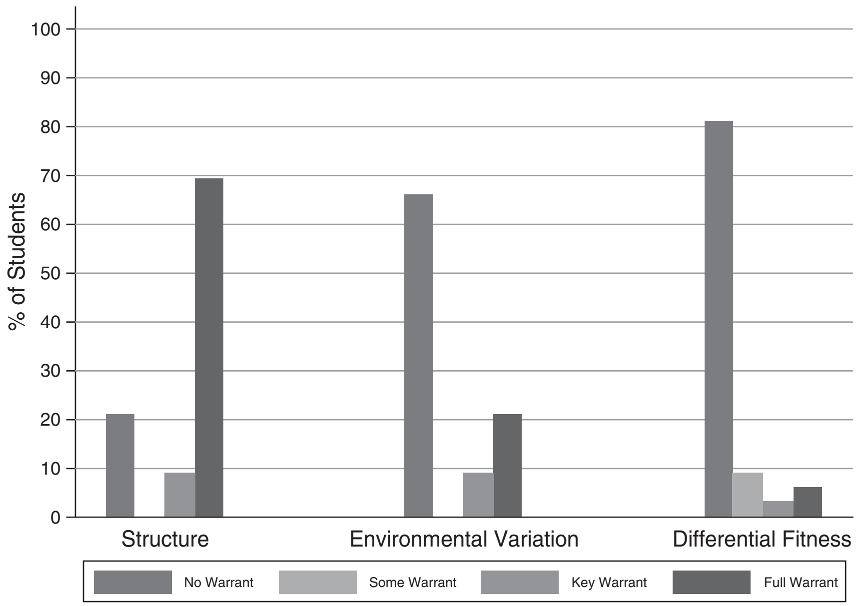

Fig. 4.2 Percentage of warranted claims in written student arguments

It could be that this problem was simply harder than our earlier problems in natural selection. Yet, this problem is structurally isomorphic to those problems. Here, students needed to identify variations in a single trait—leaf size—and relate those to variations in environmental conditions leading to differential fitness. In those other problems, students had to select from among a number of possible traits the one that produced differential fitness. On the other hand, in those problems fitness was directly shown through individual survival, whereas in this problem the idea of survival was only implied by the distribution of particular plants across a geographic range. That is, students did not see large-leafed plants die on the hot, sunny slope, they only saw that the plants that grew on that slope had small leaves. Also, if students did not really understand how the photosynthesis–transpiration compromise might affect individual plant survival, then they could not have cited it as evidence for their claims. Still, if students lacked this understanding, it also stands to reason that they would have more difficulty articulating claims of variation and differential fitness than they did.

Ideas about One's Own Warrants

A third possible explanation for the lack of supplied warrants in students' written arguments is that they may have felt that evidence was unneeded for some reason. We examined this by asking students, during a semi-structured interview, to describe why they did or did not believe their claims. In the week following the unit, individual students were pulled out of class and interviewed by the second author for about 30 minutes. In this interview, students were shown their written argument and were asked to highlight (using a marker) all of the claims they made. First, the interviewer asked students what they thought a claim was, then she provided a definition of a claim, and after that the student was asked to highlight all of their claims. When they were done, the student was asked to indicate which of their highlighted claims they were most sure of, and why. They were then asked which of the claims they were most uncertain of, and why. These two questions thus provide an idea of the standards students applied to judge certainty of their own claims—or what makes a good warrant. Finally, students were asked to describe the "best way to convince someone of something in science." This question was an attempt to find out whether or not students applied the same ideas about their own claims to what might loosely be considered scientific persuasion. We used the phrase "in science" to try to get students to think about persuading someone outside of their science class, although we cannot be sure that this is how students interpreted the phrase.

Four codes emerged from students' responses that characterized their expressed beliefs about how to warrant a claim—authority, causal, empirical, and factual (see Table 4.2). The *authority* code was used when the student warranted their claim by citing a source of authority, such as their teacher, science class, book, etc. For example, a student said she was certain of her claim, "because that is what we were talking about in science class and my teacher said it." *Causal* codes included times

Table 4.2 Codes Derived for Students' Expressed Beliefs about how to Warrant a Claim

Code	Definition
Authority	Student reason explicitly states source of authority or lack thereof for uncertain claims. Sources may include: teacher, science class, book, Internet, etc.
Causal	Student reason is based on a theoretical concept, or explanation of a theoretical concept.
Empirical	Student reason is citing some kind of empirical evidence, or lack of empirical evidence for uncertain claims. Can include: research from CENS website, data/graphs, results from experiment, etc.
Factual	Student reason is repeating their original claim by using the exact same words, paraphrasing, or rephrasing. Student asserts that their claim is a fact.

when students cited a specific scientific idea, or simply appealed to causal explanation, as in a student saying that they would convince someone in science by, "I'd explain to them why it happens and show them graphs and stuff." This statement also includes an *empirical* warrant, by appealing to graphs. Note, therefore, that single responses could possibly receive more than one code. *Factual* codes were simply assertions that the stated claim was a fact, or an explicit appeal to "facts," as in convincing someone of something in science by, " tell[ing] them a lot of facts about it."

When asked how they were certain of a particular claim, why they were uncertain of one, or how to best persuade someone "in science," most students appealed to empirical warrants. In the certain and persuasion contexts, more than half of the students cited an empirical warrant, and just under half of the students did so in the uncertain context (we note, though, that more than half of the students interviewed denied being uncertain about any of their claims). We also found that students tended to prefer a particular kind of warrant for all of these contexts. If a student cited authority as the reason for being certain of a claim, they tended to cite a lack of authority for being uncertain. We computed a preferred warrant for each student by looking to see whether or not they appealed to the same type of warrant in 2 or more of the 3 contexts. As can be seen in Fig. 4.3, empirical warrants were overwhelmingly preferred by students, $X^2(5, N = 33) = 30.21$, $p < .001$. Notice that five students had multiple preferred warrants, meaning that they gave more than one warrant in more than one context (like the example given above for causal and empirical warrants). Of these five, four expressed empirical as one of their preferred warrants.

Comparing these data to students' written arguments, we see a discrepancy. More than half of the students said that empirical warrants were how they knew they were certain of their claims, although fewer than 25% of them cited any evidence for their claims of environmental variation or differential fitness. What is going on here? There are a number of factors. First, we looked at the claims that students cited as the ones they were most certain of. We found that these claims were comprehensive, somewhat general, claims that could be induced from the data that students had looked at and their prior lab experiences in the unit. These were

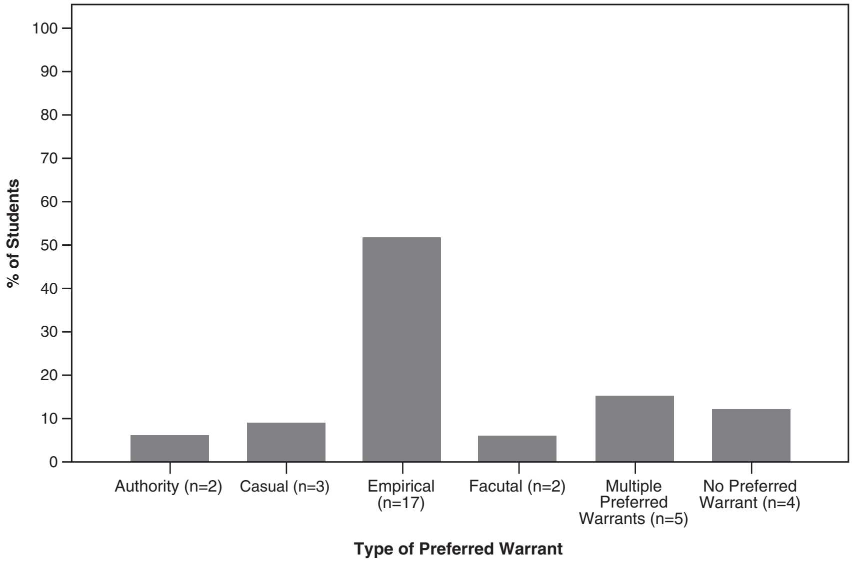

Fig. 4.3 Percentage of preferred warrants across contexts

claims like the following, "Plants are adapted to the environments that they live in," or, "plants have to balance photosynthesis and transpiration in order to survive."

Consequently, there is a mismatch between what our content analyses considered claims and what students highlighted as claims. This leads to a second type of mismatch, one of specificity. Our analyses of warrants looked for specific kinds of data, or appeals to principle, that could support specific claims related to a target conceptual framework we imposed on students' written work. Their personal warrants, their sources of belief in what they perceived as the important claims in their written work, were synthetic and generalized. That is, students appealed to the collection of data that they had looked at as providing the warrant for their general claims. On the one hand, this is a perfectly legitimate way of reasoning. On the other hand, it might be argued that it is less than ideal as scientific practice because it does not trace the specific contributions from particular data and their role in the larger story. We have to point out, however, that our findings do not permit us to say that students cannot produce such a trace, only that they do not seem to do so spontaneously, at least not in the contexts that we provided for them. (As an aside, this can explain the differences in warranting seen in students' written work here as compared to Sandoval & Millwood, 2005. In that earlier study, students were explicitly encouraged to articulate their explanations in the same terms as our coding scheme. Here, however, we applied the coding scheme post hoc to students' written work.)

We did find, however, that students who expressed a preference for empirical warrants in their interviews were more likely to provide warrants in their explanations,

$t = -2.76$, $p < .01$. Overall, students averaged 2 (out of 9) on their warrant scores, but students who preferred empirical warrants scored, on average, 4 out of 9. We take this to mean that these students have a productive epistemological belief, empirical evidence provides the best warrant for claims, that they can also apply to their own work.

Ideas about Scientists' Warrants

We queried students' ideas about how scientists warrant their claims by administering the POSE questionnaire prior to the start of the unit. We knew going in that this instrument was likely to be problematic, given the historical issues with such assessments (Kelly et al., 1998; Lederman et al., 1998). Still, we wanted to have some assessment of these students' ideas about what scientists do independent of their own scientific work, and preferred to use a validated instrument. POSE is intended by its developers to provide an overall view of students' views of the nature of science as either naïve or informed, or as in transition from naïve to informed. Our initial analyses suggested that all of our students held naïve views, so we decided to look thematically at students' responses to questions that asked them how scientists used evidence. Here we present only those data that relate to students' ideas about how scientists use evidence to support claims. These were garnered from students' responses to a general question about why scientists collect evidence or data, and a series of more specific questions about how scientists know dinosaurs really existed, how they know what dinosaurs looked like, and how it might be possible for scientists to disagree about the cause of dinosaur extinction.

Responding to how scientists know that dinosaurs existed and what they looked like, 80% of students mentioned a specific type of evidence, either "bones" or "fossils" Nearly 10% just said "evidence," and another 10% left blank responses. None of these responses was more than a word or two, and included no discussion of how "fossils" or "evidence" were used to support claims about dinosaurs.

Responding to the question of why scientists collect evidence, students gave very general responses familiar from the literature in this area. One response was of the general "to find an answer" type, given by about 40% of students. Another response was of the type "to test an idea," also given by about 40% of students. As seen by other researchers in this area (reviewed by Sandoval, 2005), such responses are quite typical for students of this age. Just over 10% of the students did not respond to this question.

Discrepancies across Contexts

In looking across these data sources, what can we say about students' epistemological beliefs about warranting scientific claims? A majority of students said that empirical evidence was how they knew (or did not know) a claim, very few of them actually provided such evidence in their own work, while an overwhelming majority of them appealed to specific forms of evidence as the reason that scientists know things. Looking more closely, there are a number of discrepancies, or at least gaps, in students' talk across their own and scientists' contexts, and in comparison to their own work.

First, there is the apparent discrepancy between students' expressed opinion that their certainty about their own claims comes largely from empirical evidence, but they were unlikely to explicitly use those data in their written explanations. We think, however, that this is not a discrepancy, or an inconsistency, between expressed belief and practice, as much as an indication of the difficulty of the practice of writing arguments. Those students with a preference for empirical warrants were significantly more likely to provide such warrants in their written work. This suggests that the preference for empirical warrants is productive for generating data-based arguments, but is not in itself sufficient to enable students to marshal appropriate warrants for claims. As we mentioned before, one issue may be that our analyses of claims and warrants occurred at a finer scale than students' own analysis during their interviews. Regardless of that, preferring evidence for your claims does not necessarily help you to find that evidence or make sense of it. Rather, it simply suggests that you are likely to mention it in your argument.

A related issue that we have not addressed is that of audience. In this intervention, the explicit audience for students' written arguments is their teacher, and includes a few other students who reviewed their work. Ultimately, even these reviews are for the teacher. Consequently, there is not much motivation to provide explicit evidence beyond the stated demand to do so as part of the task. There is no rhetorical demand for producing evidence, as it really plays no role in persuading someone of the viability of your own explanation. That is, it is quite reasonable for these students to presume that their teacher knows "the answer" to this question, and their responsibility as students is to produce that answer. While providing evidence was an explicit task demand, students may have seen it as secondary to coming to the correct conclusion.

With respect to how students talked about the evidence for their own claims compared to how they talked about scientists' evidence for claims, we see that in both contexts students understand that scientists want "evidence," and have some notions of what kind of evidence is appropriate. It appears to us that differences in knowledge about exactly that issue—what sorts of evidence are appropriate for making what sorts of claims—explains the differences in students' responses to POSE and to our interview questions. We really cannot presume that students understand anything about the scientific study of dinosaurs other than the common knowledge that scientists use fossil evidence to make claims about them. Moreover, we cannot accurately infer that students *lack* an understanding of the epistemic role of evidence, fossils in this case, in making claims about dinosaurs simply because their only response is "fossils." This is because the survey itself is ambiguous about its own rhetorical demands, an issue we return to below. Students are better able to talk about the specific evidence they have for their own claims because they are much more familiar with both the data and the claims. Thus, it may not be an epistemological difference but a domain difference. This remains an open question.

Dilemmas of Studying Epistemology

Our theory of practical epistemology (Sandoval, 2005) asserts that students' epistemological beliefs are developed through their own epistemic practices of making and evaluating knowledge claims. This view leads to the prediction that students may have different epistemological beliefs guiding their own practices of knowledge making than the ones accessed through interrogations of their opinions about scientific practice. Our findings here provide qualified support for the prediction, and thus for the theory itself. Far more students cited specific empirical evidence as the warrant for scientific claims about dinosaurs than cited empirical evidence as the reason that they knew something. Conversely, none of the student responses to how scientists know something mentioned authority, or a general causal explanation, or simply asserted the scientific claim as a fact. While a majority of students could be seen to prefer empirical warrants for themselves and for scientists, a significant minority of students appealed to other sources of warrant in their own work. This supports the common sense idea that students do not see their work in their science classes as necessarily related to what scientists do. If so, it is both an accurate conclusion, since science classes in fact rarely do resemble legitimate scientific work, and an inevitable one, as how in the world would most students get any idea what scientists actually do?

We realize that these findings provide only modest support for the theory of practical epistemologies, and raise many more questions than they answer. For us, these questions break down into questions of research methodology and instructional strategy. We focus here on the methodological issues (see Erduran, this book), as they are most pertinent to the aims of this book to outline future directions for research.

How Can We Study Epistemology?

Science educators have been studying children's ideas about science for about sixty years, and yet there is still limited understanding of how those ideas develop over time and through instruction. We know that most preadolescent students make no distinction between data and claims, and that most adolescents come to see the relation as simplistic—data show claims definitively right and wrong (see Sandoval, 2005). The issues with asking students' their views on epistemology with written survey instruments, whether they include closed or open-ended questions, are well documented (Kelly et al., 1998; Lederman et al., 1998). Interviews are similarly troublesome (Sandoval & Morrison, 2003). Such questioning seems inherently limited because of the obvious disadvantage students are at: we want to know how they understand what "real" scientists do even though we know that they have had no opportunities to see such a thing.

The alternative to asking students about what scientists do is to study what they themselves do when they learn science. Lederman et al. (1998) made this argument on methodological grounds, proposing that formal assessments would always underestimate what students know or believe. Kelly et al. (1998) made the argument both on methodological and theoretical grounds, marrying a sociocultural view of learning with a sociological view of science inspired by Latour and others to argue that a practice-oriented view of science requires analyzing students' practices of sense-making in science and comparing them to scientific practices.

The problem with a practice-only tack to studying epistemology is that looking at *what* students do is insufficient to explain *why* they do it. The approach we have described here is motivated directly by this problem: the patterns of performance in students' written arguments do not in themselves illuminate students' motives, including the goals they are pursuing and their ideas about how to satisfy those goals. What we have reported here noticeably lacks any talk of students' goals. We actually did ask students about their goals and found that those goals are, unsurprisingly, school oriented, a result consistent with the work reported by Jiménez-Aleixandre et al. (2000). Thus, our questions about certainty are instead an attempt to understand their views on how these goals are met. We find in this project, however, that we run into the same dilemma that we and others have had with more formal assessments of epistemology. The problem is that students are not very articulate about how particular pieces of data support specific claims.

Sandoval (2005) argued that studies of practice should include analyses of the artifacts that students' produce during science learning, and the discourse they engage in during the development of those artifacts. The findings we have reported here follow one of those recommendations directly—to compare properties of the artifacts that students create—in this case written arguments—and students' perceptions about how to evaluate their own artifacts. Our conclusions solidify a finding from prior research—that a majority of students prefer empirical evidence as the best warrant for a scientific claim. At the same time, we found out that a large minority of students appeal to other sources, including a notion of authority that lacks an explicit locus of that authority (i.e., the teacher is the authority because she is the teacher, not because she presents a persuasive argument herself). The other warrants offered were vague appeals to causality, and a "factual" restatement of claims that reflects what has typically been considered the conflation of claims and evidence (e.g., Kuhn, 1993a). We remain unsatisfied with these findings, however, as taking us only one step further to understanding the explicit epistemological commitments that students have. Our interview protocol, for example, failed to uncover students' explanations for why particular pieces of evidence made good warrants. One obvious next step is to develop probes that can encourage students to go beyond stating the sources of their belief (e.g., about certain or uncertain claims) to include their ideas about why those sources provide a desirable level of justification.

On the other hand, efforts to study epistemology through classroom discourse have to go further than they have to create instructional contexts in which epistemological commitments are made explicit. Our view of current research on argumentation is that these efforts, including our own, aim to put students in situations where they must make explicit epistemic decisions. That is, they have to choose the kinds of data to collect, or choose among possible interpretations of data, or choose between competing claims. All of these kinds of decisions are important decisions for students to have to make. They are central to any thorough understanding of scientific practice, and are implicated in improved student learning of scientific ideas. Still, as our study and others (reviewed by Sandoval, 2005) have shown, the need to make epistemic decisions is not accompanied by a need to make explicit the epistemological justifications that underlie those decisions. Interventions that engage students in epistemic practices, such as constructing and evaluating arguments, should also include more explicit epistemological discourse. Such a discourse would comprise the practice that is most of value studying for students' epistemological beliefs.

We remain convinced that in order to develop a more useable theory of epistemological development requires research on the epistemological ideas that students actually use while they learn and reflect on their learning—what we call practical epistemologies. Such an agenda clearly must focus on studying the learning practices that are likely to engage students' underlying epistemological beliefs, as fragmented and tacit as they may be. Argumentation is one such practice. An important direction for research on argumentation to move in is to link students' practices of argumentation with their criteria for arguments and their reasons for arguing in science class. An agenda that can span practice and ideas about practice will lead to the development of better theory, which will in turn lead to better science education.

Acknowledgments The study reported in this chapter is from the doctoral dissertation of the second author, under supervision of the first author, and was supported by a predissertation fellowship from NIMH. This work has also been supported by NSF award ESI-0352572 and through the Center for Embedded Networked Sensing (CENS), NSF CCR-0120778. The views and opinions expressed here are our own and do not necessarily reflect the views of these agencies. We are grateful to Melissa Cook, Deborah Fields, Suna Ryu, and other members of our research apprenticeship group for their helpful discussions of earlier drafts of this paper. We also thank Fouad Abd-el-Khalick for providing POSE.

References

Abd-El-Khalick, F. (2002). The development of conceptions of the nature of scientific knowledge and knowing in the middle and high school years: A cross-sectional study, Paper presented at the annual meeting of the National Association for Research in Science Teaching. New Orleans, LA.

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

Bell, P., & Linn, M. C. (2002). Beliefs about science: How does science instruction contribute? In B. K. Hofer & P. R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 321–346). Mahwah, NJ: Lawrence Erlbaum.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. Science Education, 84, 287–312.
Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. Studies in Science Education, 38, 39–72.
Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. Science Education, 88, 915–933.
Griffis, K., & Wise, J. (2005). Sensing the environment. Los Angeles, CA: Center for Embedded Networked Sensing, University of California.
Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 169–190). Mahwah, NJ: Lawrence Erlbaum.
Jiménez-Aleixandre, M. P., Bugallo Rodríguez, A., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. Science Education, 84, 757–792.
Kelly, G. J., Chen, C., & Crawford, T. (1998). Methodological considerations for studying science-in-the-making in educational settings. Research in Science Education, 28(1), 23–49.
Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. International Journal of Science Education, 20(7), 849–871.
Kelly, G. J., & Duschl, R. A. (2002). Toward a research agenda for epistemological studies in science education, Paper presented at the annual meeting of NARST 2002. New Orleans, LA.
Kelly, G. J., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. Science Education, 86, 314–342.
Kitcher, P. (1991). Persuasion. In M. Pera & W. R. Shea (Eds.), Persuading science: The art of scientific rhetoric (pp. 3–27). Canton, MA: Science History Publications.
Kuhn, D. (1993a). Connecting scientific and informal reasoning. Merrill-Palmer Quarterly, 39(1), 74–103.
Kuhn, D. (1993b). Science as argument: Implications for teaching and learning scientific thinking. Science Education, 77(3), 319–337.
Latour, B. (1987). Science in action. Cambridge, MA: Harvard University Press.
Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. Journal of Research in Science Teaching, 39(6), 497–521.
Lederman, N. G., Wade, P. D., & Bell, R. L. (1998). Assessing the nature of science: What is the nature of our assessments? Science & Education, 7, 595–615.
Lehrer, R., & Schauble, L. (2006). Scientific thinking and science literacy. In W. Damon, R. Lerner, K. A. Renninger, & I. E. Sigel (Eds.), Handbook of child psychology, 6th ed., Vol. 4: Child psychology in practice. Hoboken, NJ: Wiley.
Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. Cognition and Instruction, 11(3&4), 347–364.
Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. Journal of the Learning Sciences, 12(1), 5–51.
Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. Science Education, 89, 634–656.
Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. Cognition and Instruction, 23(1), 23–55.
Sandoval, W. A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. Journal of Research in Science Teaching, 40(4), 369–392.
Takao, A., & Kelly, G. (2003). Assessment of evidence in university students' scientific writing. Science & Education, 12(4), 341–363.
Toulmin, S. (1958). The uses of argument. Cambridge: Cambridge University Press.