Gregory J. Kelly and Richard E. Mayer, Section Coeditors

Contextual Epistemic Development in Science: A Comparison of Chemistry Students and Research Chemists

ALA SAMARAPUNGAVAN, ERIK L. WESTBY
Department of Educational Studies, College of Education, Purdue University, West Lafayette, IN 47907, USA

GEORGE M. BODNER
Department of Chemistry, College of Science, Purdue University, West Lafayette, IN 47907, USA

Received 2 May 2005; revised 27 July 2005; accepted 2 August 2005

DOI 10.1002/sce.20111
Published online 8 February 2006 in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: This study investigated the ways in which beliefs about the nature of the science vary as a function of an individual's chemistry expertise and chemistry research experience across the range from high-schools students, whose exposure to chemistry occurs in the classroom, to practicing research chemists. Interviews conducted with a total of 91 participants probed three key research questions: Do the participants' epistemic beliefs vary as a function of chemistry expertise? Are there discipline-specific values and heuristics that guide chemistry research? How does research experience influence participants' epistemic beliefs? We found that participants' epistemic beliefs varied significantly with chemistry expertise and with exposure to authentic research in chemistry. Differences in both the duration and the nature of participation in research had a significant effect on how participants conceptualized science and scientific research. We noted that only the practicing scientists saw a productive role for empirical anomalies that arise in the course of doing research. We found that research chemists thought about their scientific work in terms of a building or "engineering" model of science, rather than the classic hypothetico-deductive model of

Correspondence to: Ala Samarapungavan; e-mail: ala@purdue.edu

science invoked by some science educators. We concluded that current characterizations of the nature of science in science education may underrepresent important discipline-specific aspects of science. These results are discussed in terms of implications for science education.
 Sci Ed **90:**468–495, 2006

INTRODUCTION

A major goal of science education at all levels is to educate students about the nature of science (Abd-El-Khalick, Bell, & Lederman, 1998; American Association for the Advancement of Science, 1993; Lederman, 1986; Lederman & O'Malley, 1990; National Research Council, 1996). A challenge that we face in trying to meet this goal is the task of setting a standard against which we can measure students' answers to the question: "What is the nature of science?" We also need a better answer to the question: "Where does our knowledge about the nature of science come from?" As noted by Alters (1997), there is a broad range of philosophical positions on the nature of science and science educators typically underrepresent the range of epistemic positions in the philosophy and history of science.

Historically, the task of defining the nature of science was left to the branch of philosophy that concerned itself with the "epistemology of science." Some science educators, such as Lawson (2000), have adopted a philosophical model referred to as the hypothetico-deductive model (Popper, 1959, 1972). This philosophical model describes science as a theory-building enterprise in which systematic observation suggests hypotheses, which must be subjected to repeated experimental tests that carry the potential of invalidating or falsifying them. Theories that withstand this onslaught of repeated testing become part of the "content" of science; the scientific knowledge that is carried in our textbooks and taught in our classrooms. This model underlies many tacit assumptions within traditional science education curricula; especially the common conception of science as having two parts: "content" (the concepts, theories, and laws of science) and "process" (the use of scientific methods and reasoning).

The second half of the 20th century, however, saw the emergence of several strands of scholarship that challenged the traditional account of science described above. These new, "postpositivist" approaches vary in their details, but have some common features.

Many of the postpositivist scholars adopt a naturalistic position to questions about the epistemology of science (Giere, 1988; Goldman, 1986; Goodman, 1983; Kitcher, 1992; Quine, 1970). Proponents of this approach argue that questions about the nature of science and scientific knowledge can and should be answered empirically on the basis of data that reflect past and current practice of science. Scholars who favor a naturalistic approach often draw upon the methods and data of historical research to make empirical arguments for their epistemological theories (Kuhn, 1962, 2000; Laudan, 1984; Nersessian, 1984; Thagard, 1990). These naturalistic philosophers of science often adopt the case study method, analyzing "critical" episodes of scientific change (Conant, 1957; Kuhn, 1962; Thagard, 1990). Such case studies typically employ the analyses of archival data including the laboratory notebooks, correspondence, and publications of scientists. More recently, researchers have undertaken cognitive analyses of the research practices of working scientists (Dunbar, 1995; Sahdra & Thagard, 2003).

Naturalistic research has provided new insights into the role of interpretation in scientific decision making and the limitations of a strictly logical account of scientific inference. It has shown that scientific judgments are often based on pragmatic criteria such as the plausibility or compatibility of new ideas with those that have already been judged to be "well founded" (Laudan et al., 1986). It also suggests that epistemic norms and heuristics within different areas of science are shaped by the current conventions of disciplinary practice; for example, methodological conventions for collecting, analyzing, and presenting data. Kuhn (1977) refers to this framing role of established scientific practices in a discipline as “the disciplinary matrix.”

The recent work in history, philosophy, and cognitive studies of science suggests that there may be many kinds of science. While certain generic features of science such as attention to empirical accuracy may cut across disciplines, different areas of science develop their own contextualized epistemic norms and heuristics (Kuhn, 1977; Laudan, 1990; Longino, 1990; Thagard, 2003, 2004).

Postpositivist developments in the philosophy and history of science underlie current reform efforts in science education; especially in calls for science education curricula to attend to the nature and epistemology of science (Abd-El-Khalick & Lederman, 2000; Duschl, 1988; Hodson, 1988; Smith et al., 2000). Most of these reform efforts still focus on domain-general rather than contextual factors in epistemic knowledge and reasoning.

Recently, however, some researchers are beginning to emphasize the importance of a more contextualized domain specific approach (Duschl, 2000; Leach, Hind, & Ryder, 2003). Our current study explored contextual factors in students’ epistemic development in a specific scientific subdiscipline, that of chemistry. Drawing from naturalistic approaches to the epistemology of science described above, we are interested studying “enacted epistemologies” (Chinn & Samarapungavan, 2005). Enacted epistemologies can be thought of as functional epistemic heuristics. Members of a scientific community use shared epistemic heuristics to guide and evaluate their work and to establish bases for consensus. As noted above, recent work in the history and philosophy of science suggests that enacted epistemologies vary by disciplinary context within the sciences. Chemistry was chosen as the target domain, because prior research indicates that research chemists may employ certain unique epistemic norms and heuristics to evaluate and guide their scientific work. These heuristics appear to be embedded in a building or “engineering” model of science rather than a classic “hypothetico-deductive” theory building model (Samarapungavan, 1992; Samarapungavan, Westby, & Bodner, 1999).

The objective of the current study was to provide a description of how research chemists and chemistry students conceptualized and evaluated their work. It examined whether chemists (both practicing scientists and students) employed discipline-specific epistemic heuristics. The study also examined the influence of research apprenticeship and laboratory work on students’ epistemic development in chemistry.

In focusing on epistemologies of science, our work differs from prior research in education which takes a more global, domain general, view of epistemic development (Hofer & Pintrich, 1997; Perry, 1970; Schommer, 1990, 1993). These global approaches try to describe people’s general “theories of knowledge” without distinguishing traditions of scientific inquiry from other disciplinary traditions such as literary criticism.

Within science education, many researchers have investigated peoples’ epistemologies of science (Carey & Smith, 1993; Hodson, 1985, 1988; Hogan, 2000; Leach et al., 2000; Linn & Songer, 1993; Strike & Posner, 1985). However, this research has focused on beliefs about the nature of science in general. Often the questions asked about the nature of science are decontexualized and very abstract. For example, the Nature of Science Interview (Carey et al., 1989; Carey & Smith, 1993) includes such questions as, “what do you think the goal of science is?” and “what is a theory?” Some researchers have tried to anchor their questions about the nature of science in concrete examples. One of the questions on the Views of Nature of Science Questionnaire—Form C is—“Science textbooks often define a species as a group of organisms that share similar characteristics and can interbreed with one another to produce fertile offspring. How certain are scientists about their characterization of what a species is? What evidence do you think scientists use to determine what a species is?” (Lederman et al., 2002). However, even when questions about the nature of science are anchored with such concrete examples, sophisticated epistemic answers require considerable domain specific knowledge.

Our study takes a somewhat different approach by anchoring epistemic questions in our participants' current work within a specific subdomain of science, chemistry. By doing so, we hope to explore participants' enacted epistemologies and how these vary with increased research expertise. Previous research indicates that high-school students' epistemologies of science are both simple and decontexualized. Carey and Smith (1993) suggest that students do not have a sense of either the provisional nature of scientific knowledge or the cyclical and constructive nature of scientific research. They are, for instance, unable to articulate epistemic heuristics that might help a scientist follow up on experimental failures. Work of this nature led us to expect that epistemologies of science might undergo considerable change as students move from high school through graduate training to become practicing scientists.

Description of the Study

This study investigated the ways in which beliefs about the nature of the science vary as a function of an individual's chemistry expertise and chemistry research experience. The data for this paper are drawn from interviews with research chemists (Samarapungavan, 1992; Westby, Samarapungavan, & Bodner, 2000) and interviews with high-school, undergraduate, and graduate chemistry students (Westby, 2003). The interviews were based on sets of questions that examined: (a) the contextualized criteria that the participants use to evaluate and regulate their own work in chemistry and (b) the participants' general views about what science is. We hypothesized that the questions anchored in the participants' chemistry work were more likely to elicit discipline-specific epistemic concepts. The interviews are described in more detail in the methodology section. The interviews were designed to provide data with regard to three key research questions:

- Do the participants' epistemic beliefs vary as a function of chemistry expertise?
- Are there discipline-specific values and heuristics that guide chemistry research?
- How does research experience influence the participants' epistemic beliefs?

METHODOLOGY

The methodological framework for this study draws from cognitive science methods for the gathering and analysis of verbal data (Chi, 1997). Specifically, data were collected through scripted, semistructured interviews. Iterative quantitative content analysis was used to code and analyze verbal data (Samarapungavan & Wiers, 1997; Vosniadou & Brewer, 1992).

Sample

The sample consisted of 91 volunteers selected from five groups that represent different degrees of experience with chemistry and chemical research. A brief description of each group follows: The first group ("scientists") were chemistry faculty actively involved in chemistry research ($N = 13$); the second group ("grads") were chemistry graduate students enrolled in a doctoral program ($N = 22$); the third group ("research trainees") were undergraduate chemistry majors involved in chemistry research who were considered to be research trainees ($N = 20$); the fourth group ("undergrads") were second-semester undergraduates enrolled in a general chemistry course with a laboratory component ($N = 17$); the participants in the four groups described above were from a major midwestern research university. The fifth group ("high school") were students at a large midwestern public high school who were enrolled in a high-school chemistry course with a laboratory component ($N = 19$).

Purposeful sampling, based on peer nomination, was used to identify the participating scientists and graduate students (Patton, 1990). Members of the other three groups were volunteers, which meant that our sample is only representative of students who have a special interest in science.

Epistemic Interviews

Content and Structure of Interviews. The study from which the current data are drawn examined participants' beliefs with to several epistemic themes using a scripted, semistructured interview format. The epistemic themes were identified by members of our research team through a theoretical analysis of the literature in the naturalistic history and philosophy of science as well as a review of prior work in science education cited above. This paper analyzes and reports data focused on five epistemic themes which will be described below. The interviews were based on a set of sequentially ordered core questions, with follow-up questions asked for clarification where needed. The interview questions used in this study were adapted from an earlier interview with scientists (Samarapungavan, 1992). To help ensure content validity, the initial interview questions were reviewed and critiqued by two philosophers with interests in the history and philosophy of science who were faculty members at major research universities. Based on these reviews and pilot data, the questionnaire was modified and a set of different, partially overlapping interviews were developed for each of the groups (Westby, 2003).

The specific questions analyzed in this study are listed in Table 1. Some questions focused on discipline-specific aspects of epistemology (Q1–Q11), whereas others examined the participants' beliefs about science in general (Q12, Q13). All interviews covered a common core of epistemic themes, but the specific questions were adapted for different levels of expertise (see Table 1).

The questions in Table 1 are organized into a set of five epistemic themes. *Theme 1: Description of Own Work* contains questions that allowed the participants to describe their current work as chemists or chemistry students. Scientists, graduate students, and research trainees were asked to describe their current research and to explain how they came to be involved in this research, while undergraduate and high-school students enrolled in chemistry laboratory courses were asked to describe their work in their respective courses and to explain the purpose of their laboratory work. *Theme 2: Choice of Problems and Methods* consists of questions that addressed the choice of problems and methods. Scientists, graduate students, and research trainees were asked how they come to choose the specific problems and methods they employ, whereas undergraduate and high-school students were asked what determines the specific methods that they use in their laboratory courses. *Theme 3: Models for Handling Empirical Anomalies* consists of questions that probed the participants' views of empirical anomalies in laboratory work. The participants were asked to describe their strategies for handling empirical anomalies. Scientists, graduate students, and research trainees were also asked how they made decisions about whether to continue with a specific set of ideas. *Theme 4: Criteria for Evaluating Own Work* consists of questions that probed the participants' criteria for evaluating their own chemistry work. *Theme 5: What Is Science* consists of questions that examined the participants' concepts of science and to describe what differentiates the sciences from disciplines that are not sciences.

TABLE 1
Epistemic Interviews: Questions by Theme

Theme 1. Description of Own Work

Groups 1, 2, and 3:
Q1. Could you tell me about the work you are doing this year?
Groups 4 and 5:
Q1. Could you tell me about the work you are doing in chemistry lab?
Q2. What is the purpose of labs?

Theme 2. Choice of Problems and Methods

Groups 1, 2, and 3:
Q2. How did you become involved in this research?
Q3. How did you come to choose the specific problems that you are currently working on?
Q4. How did you decide upon specific methods to investigate these problems?
Groups 4 and 5:
Q3. N/A
Q4. What is the purpose of the specific methods that you use in labs?

Theme 3. Models for Handling Empirical Anomalies

Groups 1, 2, and 3:
Q5. In the course of your research, have you performed experiments that have yielded unexpected results? (If yes) Can you give me some examples of such situations?
Q6. (Follow-up). How do you decide what to do next? How do you evaluate outcomes?
Q7. In the course of your research career have you ever had occasion to give up your ideas?
Q8 (Follow-up). How do you decide whether to continue with an idea or to give it up?
Groups 4 and 5:
Q5. In the course of your labs have you obtained unexpected results? (If yes) Can you give me some examples of such situations?
Q6. (Follow-up 5). How do you decide what to do next? How do you evaluate outcomes?

Theme 4. Criteria for Evaluating Own Work

Groups 1, 2, and 3:
Q10. How do you evaluate the success of your work?
Q11. What do you consider to be your best work? Why do you think it is your best work?
Groups 4 and 5:
Q10. How do evaluate the success of your work in chemistry?
Q11. (Same as for groups 1, 2, and 3)

Theme 5. What Is Science? (The same questions were used for all groups.)

Q12. What is science?
Q13. What differentiates the sciences from disciplines that are not sciences?

Interview Procedure. Each participant was interviewed individually for approximately 50–60 min. The interviews were audiotaped and transcribed. A member of the chemistry research faculty served as a domain expert and was present during all interviews with the scientists. This member of the research team was able to clarify and explain technical aspects of the participants' answers (e.g., methodological descriptions) for subsequent coding and analysis.

Coding and Analysis of Data

Quantitative content analysis was used to analyze the data. The coding procedures were adapted for the current domain from prior research (Chi, 1997; Vosniadou & Brewer, 1992).

Responses for each participant were analyzed by theme. Responses to each set of questions corresponding to a specific theme were analyzed independently of all other responses using a two-stage coding procedure. We grouped the participants' responses into nominal categories based on common meaning or ideas. Each nominal category was identified by a unique code (N1, N2, etc.). The nominal coding categories derived from the raw data were then analyzed and grouped into ordinal levels (O1, O2, etc.). Lower ordinals codes represent higher or more sophisticated levels of response. The coding categories for each theme are shown in Table 2. In these tables, the coding categories are organized hierarchically, with nominal codes grouped under their corresponding ordinal codes. For each theme, each participant's response receives only one nominal and one ordinal code. The coding process is described in detail below.

Nominal Coding. During this initial stage of coding, we used a bottom-up scoring procedure, in which we tried to identify and code for common meaning. The participants' responses to our questions were often complex and varied in the range of ideas expressed. Consequently, there was sometimes a partial overlap in the content of responses coded into separate nominal categories. To reflect the range of variation in the specific ideas expressed, each response that contained unique meaning or content not included in existing nominal categories was assigned to a new nominal category. A response was assigned to a nominal coding category only if it contained all the key elements for that category. The key elements for each nominal code by theme are provided in Table 2.

As an example of nominal coding, consider responses to questions in Theme 1 (Description of Own Work). The participants assigned to nominal code "N1" described their current work in terms of creating new materials (as shown in Table 2). These participants also said that they chose their own problems and methods, and they articulated epistemic reasons for doing so. An example follows:

> One thing we are working on right now is a fairly new material called full-luminescent porous silicon . . . [which] can emit light under radiation and this is really important for a number of reasons . . . to build new types computers based on light signals instead of electrical signals . . . the surface is covered with silicon-hydride groups, so we are using these as chemical handles with which we can functionalize that surface, because we chemists look at this surface and say, "Oh, silicon-hydride groups, I know what to do with those." Whereas really, this material has been in the hands of physicists who have no idea, and they just blast it with ion beams and smear plastic all over it, very sort of brutal things like this . . . (Scientist 1)

The participants assigned to nominal code "N4" also described their current work in terms of creating new materials. However, they said that the choice of research problems and methods was determined for them by someone else, such as a more senior member of their research group or their graduate advisor (see Table 2). An example follows:

> Since two summers ago, I work with [name deleted][1] doing asymmetric catalysis with olefins using a catalyst that one of the graduate students designed . . . I work with one of the graduate students, it's his project and I am basically doing whatever he wants the next step to be. . . (Research Trainee 1)

This bottom-up nominal coding procedure allowed us to preserve conceptual variation in responses while grouping common responses together. Once all item-level data were

[1] All proper names have been deleted from quotations to preserve confidentiality.

TABLE 2
Ordinal and Nominal Response Codes by Theme

Theme 1: Description of Own Work

O1-N1. Making new or better materials (synthesizing molecules etc.) + explanation in terms of epistemic or pragmatic factors (e.g., yield, elegance, novelty, utility)
O1-N2. Developing new instruments/processes for analyzing materials etc. (e.g., spectroscopy techniques) + explanation in terms of epistemic or pragmatic factors
O1-N3. Understanding systems, /properties of materials, etc., + explanation in terms of epistemic or pragmatic factors
O2-N4. Making new or better materials + part of larger research group effort directed by advisor/boss
O2-N5. Developing new instruments/processes etc. (e.g., spectroscopy techniques) + part of larger research group effort directed by advisor/boss
O2-N6. Understanding systems, /properties of materials, etc., + part of larger research group effort directed by advisor/boss
O3-N7. Course content; refers to chapter, lab exercise etc.

Theme 2: Choice of Problems and Methods

O1-N1. Interplay of multiple factors: (a) interest, (b) desire to do good/novel science, (c) utility/practical value, (d) prior training, (e) resources (money, students etc.)
O2-N2. Guidance from lab director + use standard methods of research group or discipline
O3-N3. Lab director or other senior research group member assigns specific tasks
O4-N4. Course requirement/instructor + purpose is to learn standard techniques, instrumentation, /demonstrate key concepts
O4-N5. Course requirement/instructor + purpose is to keep things fresh, prevent boredom, learn to be organized and follow directions
O4-N6. Course requirement/instructor + purpose is to prepare for college chemistry
O4-N7. Course content + no purpose − labs do not make sense

Theme 3: Models for Handling Empirical Anomalies

O1-N1. Distinguish between routine (technical) and productive anomalies that lead to new insight + build ways to distinguish between them in methodology
O2-N2. Anomalies are routine technical errors: Identify/fix error, try different technique, read literature and seek advice of other experts
O3-N3. Anomalies are routine technical errors: redo everything from scratch
O4-N4. Initial response has content focus (content is hard/boring etc.) + attribute empirical anomalies to personal "stupidity," lack of skill. No strategies for dealing with anomalies
O4-N5. Initial response has content focus + get correct answer from friends/fudge data
O4-N6. Initial response has content focus + follow up for labs: no problems (labs are easy)

Theme 4: Criteria for Evaluating Own Work

O1-N1. Multiple criteria: (a) pragmatic: peer recognition, fruitfulness, relevance to society, (b) design: elegance, efficiency, yield, (c) cognitive: novelty, difficulty, completeness of work
O2-N2. Cognitive criteria: solving a novel/difficult problem, completeness of work
O3-N3. Get results expected/able to "run" procedure accurately
O3-N4. Research lab supervisor or boss is happy
O4-N5. Complete required class assignments/good grades/ positive feedback from instructor
O4-N6. Personal interest—the topic is interesting to me
O4-N7. Do not know/do not like any of my current work

Theme 5: What Is Science?

O1-N1. Understand natural world through scientific methods + objectivity based on standard procedures for gathering evidence + mention specific norms/standards for evaluation
O2-N2. Understand natural world + based on objective evidence
O2-N3. Understand natural world + focus on cognitive dimensions: explanation, analysis etc.
O2-N4. Making sense of the world at a personal level + relativist stance on knowledge
O3-N5. Utilitarian aspects only (solve practical problems, improve human condition)
O3-N6. Study of the natural world—refer to content (e.g., atoms and molecules)

assigned a nominal code, an external rater was asked to independently score one third of the interviews from each group in the study (selected randomly). Inter-rater agreement for nominal coding was 98%. Members of the research team reviewed the responses on which there was disagreement and final codes were assigned through discussion.

Ordinal Coding. Once the nominal coding was completed, the nominal codes for each theme were further grouped together into ordinal categories (O1, O2, etc.). A lower numerical code, such as O1, corresponds to a more advanced or sophisticated response than a higher code such as O2. There are three levels of ordinal coding (O1–O3) for Theme 1 (Description of Own Work) in Table 2, for example. O1 includes nominal codes N1–N3, O2 includes nominal codes N4–N6, and O3 includes nominal code N7. The ordinal coding of nominal categories was based on correspondence to epistemic concepts articulated in the history and philosophy of science, as well as on the range, elaboration, and specificity of the ideas expressed. This is illustrated by the ordinal coding of responses for Theme 4 (Criteria for Evaluating Own Work), which explored the participants' criteria for evaluating their own scientific work, shown in Table 2. O1 responses for this theme expressed the use of multiple evaluation criteria including design criteria (e.g., yield, selectivity), pragmatic criteria such (e.g., impact, fruitfulness, relevance to social problems), and intellectual or rational criteria (e.g., novelty, completeness of work). O2 responses were narrower in scope, referring only to intellectual or rational criteria. O3 responses reflected a "job" orientation to laboratory research, in which the ability to "get the job done" in a timely manner, or the satisfaction of the research supervisor, were the primary criteria for evaluating success. O4 responses in this theme did not refer to any specific aspects of the laboratory work at all. Instead these participants used very general, unelaborated heuristics such as getting good grades on course work, or whether the topic was "boring." Inter-rater reliability was computed independently for ordinal coding using the same procedure as for nominal coding. The inter-rater agreement was 92%. Disagreements were resolved through discussion.

Analysis. After the ordinal coding was completed, nonparametric (Kruskal–Wallis) tests were performed for each theme to examine if the ordinal levels of response varied significantly across the five groups included in the study. The Kruskal–Wallis test is a nonparametric alternative to the ANOVA for multiple independent samples (Hollander & Wolfe, 1999; Kruskal & Wallis, 1952). It is appropriate when important assumptions for parametric statistics such as an interval level of measurement, normal distribution of the variable in the source populations, and equal variance in the samples cannot be met (Hollander & Wolfe, 1999). An α level of $p = .001$ was selected as the criterion for statistical significance. If the Kruskal–Wallis analysis indicated that there were significant differences between groups, post hoc multiple comparison procedures based on pairwise rankings were used to further explore group differences (Hollander & Wolfe, 1999; Steele, 1960). An α level of $p = .05$ was selected as the criterion for statistical significance for the post hoc comparisons. The results of these tests and descriptive data are presented below.

RESULTS

Results of Kruskal–Wallis

The results of the Kruskal–Wallis tests indicate there were significant differences in group responses for each of the five themes (see Table 3). To further explore the nature of these differences, we conducted post hoc pairwise comparisons of group performance for each theme. The results of the post hoc comparisons and descriptive data for each theme are presented below.

TABLE 3
Kruskal–Wallis Test of Significance for Thematic Response Level by Group

Theme	*df*	Chi Square
1: Description of Own Work	4	79.754*
2: Choice of Problems and Methods	4	87.839*
3: Models for Handling Empirical Anomalies	4	85.178*
4: Criteria for Evaluating Own Work	4	76.014*
5: What Is Science?	4	23.064*

Note. *denotes significance at $\alpha = .001$ level.

Theme 1: Description of Own Work. Responses to questions in Theme 1 were obtained by asking each group of participants to describe the work that they were engaged in over the past year. Post hoc pairwise comparisons by group (see Table 4) indicated that the scientists differed significantly in their responses from all other groups. The graduate students and undergraduate research trainees did not differ significantly from each other, but each of these groups differed significantly from the undergraduates and high-school students. There were no significant differences between undergraduate and high-school students.

Table 5 shows the percentage of ordinal responses for each group. All of the scientists, 82% of the graduate students, and 40% of the undergraduate research trainees, gave O1 responses. The category of O1 responses included three variants at the nominal level. N1 respondents described their work as that of building molecules, N2 respondents described their work as developing new instruments or techniques for chemical analysis, and N3 respondents described their work as trying to understand the behavior or properties of chemicals. All the participants with O1 responses spontaneously explained their projects in terms of certain pragmatic and/or epistemic goals. Specific examples of O1 responses follow:

Example 1 (O1-N1)

> I am a synthetic organic chemist and what that means is that we are building molecules . . . our interest in these molecules, I guess, is threefold. First of all we would like to be able to make them in an efficient way at the strategic level. Two, we would like to be able to make them in an efficient way at the tactical level, meaning the best possible overall yield. And three . . . we would like to answer biological questions . . . virtually all of our target molecules have major biological interest in the terms of cancer, heart disease . . . and in addition to that, we want new chemistry to evolve . . . (Scientist 4)

Example 2 (O1-N2)

> It is called combinatorial chemistry. So this technology allows science to grow much faster . . . I spent my thesis research preparing artificial receptors and it took me an average of six months to make one single receptor. Using combinatorial chemistry, I can make one million or one billion receptors in a few months . . . (Scientist 3)

Example 3 (O1-N3)

> We try to understand the behavior of chain molecules when they are in the close vicinity of a surface or an interface . . . " (Scientist 7)

TABLE 4
Standardized Difference Scores for Post Hoc Pairwise Comparisons by Group and Theme

	Grads	Research Trainees	Undergrads	High School
Theme 1: Description of Own Work				
Scientists	.541	5.390*	5.498*	5.627*
Grads	–	1.409	5.686*	5.863*
Research Trainees	–	–	4.246*	4.373*
Undergrads	–	–	–	0.000
Theme 2: Choice of Problems and Methods				
Scientists	1.490	3.842*	6.592*	6.748*
Grads	–	2.743*	5.907*	6.091*
Research Trainees	–	–	3.214*	3.310*
Undergrads	–	–	–	0.000
Theme 3: Models for Handling Empirical Anomalies				
Scientists	2.474*	3.318*	6.782*	6.942*
Grads	–	1.025	5.057*	5.215*
Research Trainees	–	–	3.991*	4.110*
Undergrads	–	–	–	0.000
Theme 4: Criteria for Evaluating Own Work				
Scientists	2.354*	3.608*	6.601*	6.942*
Grads	–	1.495	4.981*	5.349*
Research Trainees	–	–	3.476*	3.788*
Undergrads	–	–	–	0.200
Theme 5: What Is Science?				
Scientists	2.694*	2.748*	3.877*	3.014*
Grads	–	1.118	1.506	0.455
Research Trainees	–	–	1.363	0.330
Undergrads	–	–	–	1.030

Note. *denotes significance at $\alpha = .05$ level.

O2 responses were given by 18% of graduate students and 60% of research trainees. These participants typically talked about their specific work as part of a larger research group effort directed by others.

Example 1 (O2-N6)

> As a group, we are working on identifying the structure of different kinds of proteins as they move from liquid to gas or from a solid to a liquid . . . I am just actually

TABLE 5
Frequency and Percentage of Nominal and Ordinal Responses by Group and Theme

	Scientists		Grads		Research Trainees		Undergrads		High School	
Codes	Fr.	(%)	Fr.	(%)	Fr.	(%)	Fr.	(%)	Fr.	(%)
Theme 1: Description of Own Work										
O1-N1	4	(31)	3	(14)	6	(30)	0	(0)	0	(0)
O1-N2	4	(31)	5	(23)	1	(5)	0	(0)	0	(0)
O1-N3	5	(38)	10	(45)	1	(5)	0	(0)	0	(0)
O1 Total	13	(100)	18	(82)	8	(40)	0	(0)	0	(0)
O2-4	0	(0)	2	(9)	2	(15)	0	(0)	0	(0)
O2-5	0	(0)	0	(0)	3	(10)	0	(0)	0	(0)
O2-6	0	(0)	2	(9)	7	(35)	0	(0)	0	(0)
O2 Total	0	(0)	4	(18)	12	(60)	0	(0)	0	(0)
O3−N7*	0	(0)	0	(0)	0	(0)	17	(100)	19	(100)
Theme 2: Choice of Problems and Methods										
O1-N1*	12	(92)	3	(14)	0	(0)	0	(0)	0	(0)
O2-N2*	1	(8)	19	(86)	0	(0)	0	(0)	0	(0)
O3-N3*	0	(0)	0	(0)	20	(100)	0	(0)	0	(0)
O4-N4	0	(0)	0	(0)	0	(0)	17	(100)	8	(43)
O4-N5	0	(0)	0	(0)	0	(0)	0	(0)	5	(26)
O4-N6	0	(0)	0	(0)	0	(0)	0	(0)	5	(26)
O4-N7	0	(0)	0	(0)	0	(0)	0	(0)	1	(5)
O4 Total	0	(0)	0	(0)	0	(0)	17	(100)	19	(100)
Theme 3: Models for Handling Empirical Anomalies										
O1-N1*	13	(100)	1	(5)	0	(0)	0	(0)	0	(0)
O2-N2*	0	(0)	20	(90)	10	(50)	0	(0)	0	(0)
O3-N3*	0	(0)	1	(5)	10	(50)	0	(0)	0	(0)
O4-N4	0	(0)	0	(0)	0	(0)	3	(18)	0	(0)
O4-N5	0	(0)	0	(0)	0	(0)	4	(23)	2	(10)
O4-N6	0	(0)	0	(0)	0	(0)	10	(59)	17	(90)
O4 Total	0	(0)	0	(0)	0	(0)	17	(100)	19	(100)
Theme 4: Criteria for Evaluating Own Work										
O1-N1*	13	(100)	1	(5)	0	(0)	0	(0)	0	(0)
O2-N2*	0	(0)	15	(68)	2	(10)	0	(0)	0	(0)
O3-N3	0	(0)	4	(18)	14	(70)	1	(6)	0	(0)
O3-N4	0	(0)	1	(5)	4	(20)	0	(0)	0	(0)
O3 Total	0	(0)	5	(22)	18	(90)	1	(6)	0	(0)
O4-N5	0	(0)	0	(0)	0	(0)	14	(82)	19	(100)
O4-N6	0	(0)	0	(0)	0	(0)	2	(12)	0	(0)
O4-N7	0	(0)	1	(5)	0	(0)	0	(0)	0	(0)
O4 Total	0	(0)	1	(5)	0	(0)	16	(94)	19	(100)
Theme 5: What Is Science?										
O1-N1*	6	(46)	0	(0)	0	(0)	0	(0)	0	(0)
O2-N2	6	(46)	16	(72)	12	(60)	7	(41)	13	(68)
O2-N3	1	(8)	0	(0)	3	(15)	1	(6)	1	(5)
O2-N4	0	(0)	1	(5)	0	(0)	0	(0)	0	(0)
O2 Total	7	(54)	17	(77)	15	(75)	8	(47)	14	(72)
O3-N5	0	(0)	2	(9)	1	(5)	1	(6)	1	(5)
O3-N6	0	(0)	3	(14)	4	(20)	8	(47)	4	(21)
O3 Total	0	(0)	5	(23)	5	(25)	9	(53)	5	(26)

Note. *denotes that there is no separate row for the ordinal total because there is only one nominal code for the corresponding ordinal code.

> repeating stuff that was done before, so I am just following a procedure multiple times to see if it's sound. (Research Trainee 6)

All of the undergraduate and high-school students enrolled in chemistry laboratory courses gave O3 responses. These responses typically described current work with reference to course work requirements. There was no attempt to frame laboratory work in terms of its scientific function or purpose.

Example 1 (O3-N7)

> I don't know . . . It [lab work] doesn't make much sense, when you do it a lot of it, you don't know why you are doing it. The spectrophotometer, with the frequencies of light, we took the zinc out of a penny and only left a shell. There was one more but I can't remember it. (High School 7)

Theme 2: Choice of Problems and Methods. Theme 2 questions probed the participants' ideas about the factors that determined the problems they were currently working on in the laboratory and the methods that they used to tackle these problems. For scientists, graduate students, and research trainees, Theme 2 questions were anchored in their current research. For undergraduate and high-school students, the questions were anchored in their current laboratory work.

The results of the post hoc comparisons indicate that there were no significant differences in responses between scientists and graduate students and between undergraduate and high-school students (see Table 4). All other groups differed significantly from each other.

O1 responses were given mainly by scientists (92%). The participants in this category gave multifaceted responses, listing a combination of pragmatic and epistemic factors that led them to select certain research problems and methods for tackling those problems.

Example (O1-N1)

> My primary goals in choosing research projects? Number one [that] it be what I call good science, which means that you are attacking a fundamental question in some quantitative fashion, hopefully with a yes/no answer . . . Really, in most cases, that doesn't turn out to be true, that is a basic goal. And that it match my students' areas of interest and long-term goals. So the research is very much a teaching vehicle for the students involved . . . if it is solvable by [technique[2]], it is more likely to get my attention than if it is not . . . (Scientist 5)

O2 were provided primarily by graduate students (86%) although one practicing scientist gave an O2 response as well (see Table 5). Responses in this category attributed the choice of research problem to guidance from a laboratory director or doctoral advisor. Additionally, the participants in this category said that they were using standard methods in their discipline.

Example (O2-N2)

> The general idea was given to us by [name deleted] . . . you only get a skeleton of the project and you have to fill in . . . mostly through literature searches and talking to people about what has worked for them . . . (Grad 6)

[2] Names of specific techniques are deleted to preserve confidentiality.

All the research trainees gave O3 responses, while none of the other groups did (see Table 5). O3 responses focused on specific laboratory tasks rather than the research project as a whole and attributed the assignment of the tasks to the director of the research group or other senior members.

Example (O3-N3)

> In terms of the specific molecules I am looking at? That comes from the director; he lets me know what to work with and what to look for . . . (Research Trainee 5)

All the undergraduate and high-school students gave O4 responses (see Table 5). These participants said that the laboratory work was determined by course requirements and/or the course instructor. Such responses were consistent with the structure and demands of typical college and high-school laboratory courses. An interesting aspect of the O4 responses was the participants' beliefs about the purpose of the laboratory work that they were doing. Within the O4 category, three nominal response types were differentiated based on the purposes ascribed to laboratory work (see N4, N5, and N6 in Table 2). N4 respondents were primarily undergraduate students who said that the purpose of the laboratory work was to illustrate key chemistry concepts and/or to teach them standard techniques of chemical analysis and the use of instrumentation. N5 respondents were high-school students who said that the purpose of laboratory work was to prepare them for college chemistry courses. N6 respondents were high-school students who said that the purpose of laboratory work was to provide a break from the regular school routine, to prevent boredom, or to see if students were organized and could follow directions. Specific examples follow:

Example 1 (O4-N5)

> . . . Mr. [name deleted] has said that a lot of college science classes have laboratory sections, and that they have to do lab reports too so maybe it's to like get us familiar with it. (High School 13)

Example 2 (O4-N6):

> Well the course covers chemical equations and we are doing a lot of stoichiometry and mole conversions . . . [Follow-up] to like see if we can follow directions and do what we are supposed to do. (High School 6)

Theme 3: Models for Handling Empirical Anomalies. As noted earlier, attempts to characterize the empirical dimensions of science have been at the heart of scholarship and debate about the nature of science. The questions for this theme probed the participants' views of the empirical aspects of scientific work, especially their views about the nature and sources of empirical anomalies. There were also questions that examined the specific strategies that the participants used to handle the empirical anomalies.

The post hoc comparisons indicate that there were significant differences in responses on all pairwise comparisons except between graduate students and research trainees and between undergraduate and high-school students, respectively (see Table 4). All the scientists gave O1 responses (see Table 5). These participants said that empirical anomalies were frequent and pervasive in research. They also distinguished between two types of empirical anomalies. The most frequent source of empirical anomalies was routine methodological error. O1 participants identified many chemistry specific sources of routine error such as temperature fluctuations, improper calibration of instruments, cross-contamination or impurities in chemicals, and the technical expertise of laboratory assistants and students ("good hands"). These participants said that they built checks and controls for routine errors into their research, by anticipating the kinds of errors that might occur, planning experimental replications to establish the robustness of their findings, and planning sequences of experiments to triangulate and understand their empirical results. However, O1 participants also identified what we call "productive" anomalies. These are empirical results that are unexpected and (initially) problematic but the process of trying to understand and account for the anomalous findings leads to exciting new insights or venues for research. Below is an example of an O1 response that illustrates the distinction between routine and productive anomalies:

Example (O1-N1)

> [it] deals with more intangible things such as being a chemical detective and an opportunist, in that many times you write down on a piece of paper what a reaction is supposed to do and it does something quite different and there is a great deal of fun in trying to figure out what it did wrong . . . many times when one is attentive and puts these things in the right perspective, you discover that nature is giving you something much more than you asked for in the first place . . . You know we have a lot of computer modeling that we do before we run the reactions that tell us at least in general detail what we expect and if something very different happens, that means we haven't considered something that nature thinks is a lower energy process than we did, and nature never makes a mistake, only chemists make mistakes . . . [We have given up research projects] many times, we would be trying to synthesize a given molecule and we would run out of good ideas to try . . . In organic chemistry the worst decision made, a faculty member has to make, is when is it time to say enough; because in the back of your mind, if the work is not done with your own hands, and it has been many years since the work has been done with my own hands, that you kind of wonder that the reason it is failing is because the student hasn't quite mixed the things right, put in the right catalyst, weighed the right amounts, and followed the reaction in the right way . . . Students are a great combination of things, some are good in the library, some are good very much in the lab, the so called golden hands approach . . .
>
> . . . Well, in the last couple of years we have had some [empirical anomalies] which are sufficiently interesting that they have changed my research direction in very large ways. Three years ago, we made a discovery which was the inception of making a carbon–carbon bond from a carbon–hydrogen bond. Basically, it hadn't really been done before and we discovered this much in the way of trying to dissolve a material in a solvent and run a reaction and the material reacted with the solvent and from there it directed our research for the next year and a half and had me put three or four new people on such a project . . . The other thing is, my projects have a lot of interrelated phenomenon, that somebody in lab A makes a discovery and gets a failure and somebody in lab D has been trying to do [it] a different way and this sort of cross fertilization approach is very, very useful, and that is why I try to have many projects that are not identical but they all use similar sorts of strategies and we find things. (Scientist 4)

Additionally, O1 respondents said that they had abandoned specific research projects and redirected their attention to others based on a variety of pragmatic factors. Some of the factors mentioned were laboratory resources such as equipment and technical expertise of lab personnel, funding, the estimated time required to solve the target problem relative to the time it would take to succeed on other projects within the lab, and the time it would take competing labs in the scientific community to solve the same problem. However, O1 respondents also said that they never really abandoned a research idea (unless another research group beat them to the solution); rather, they put less successful or more difficult projects on hold until some future date when they hoped to have more resources. An example follows:

Example (O1-N1)

> . . . I guess in your heart of hearts, you never really give [a research project] up. You are just waiting for the right person to come along. Reformulation, oh, yeah, that happens all the time . . . (Scientist 2)

O2 responses were given by 90% of graduate students and 50% of undergraduate research trainees. Like O1 respondents, those participants who gave O2 responses recognized that empirical anomalies were common. They could also identify common sources of routine methodological error and used a variety of domain specific trouble shooting techniques to try and locate and correct for such error. However, unlike O1 respondents, they viewed empirical anomalies negatively, as something to be eliminated and they did not talk about productive anomalies.

Example (O2-N2)

> Yes, that [empirical anomaly] is just a normal part of science. . . one time when I had confusing results, my guess is that the amount of protein I was using was too high or something, so the first thing that I looked at was what parameters was I using . . . I think the problems that I mostly encounter were related to the techniques. (Grad 2)

O3 responses occurred primarily among research trainees (50%). These respondents viewed anomalies as their personal failures to produce what was needed for the research project. O3 responses are characterized by a low-level task or procedural focus and a lack of discipline-specific heuristics for trouble shooting. Typically, these participants responded to empirical anomalies by redoing everything from scratch.

Example (O3-N3)

> . . . My biggest problems in the lab have to do with making the product and not having all of the imperfections . . . Try it over, my first step would be to do it over exactly the same . . . mostly it's just trial and error, redo, redo process. (Research Trainee 5)

All the undergraduate students and high-school students gave O4 responses (see Table 5). O4 participants typically responded to the initial questions by referring to the difficulty of chemistry concepts rather than to empirical anomalies. Upon follow-up questioning to redirect their attention to empirical anomalies, they responded with one of three variations at the nominal level (see N4, N5, and N6 in Table 2). The participants coded N4, seemed to regard anomalies as infrequent occurrences and attributed the odd instance of laboratory error to their lack of personal skill or knowledge or to careless mistakes. They did not seem to have any specific strategies for dealing with empirical anomalies. N5 participants said that they dealt with anomalies by copying the correct answers from friends or fudging their data to get the expected results. N6 participants denied experiencing any empirical problems in the course of laboratory work. This was the most common variant of O4 responses with 59% of undergraduates, and 90% of high-school students saying they never had any problems in the course of their lab work! Specific examples follow:

Example 1 (O4-N5):

> I guess the point of lab is to show the practical application but a lot of time, it's like we just have two hours to get done, so we just hurry up and do what we are supposed to do . . . we just want to get the right percent yield so we get a good grade . . . We kind of know what were are supposed to get before we do the lab . . . so we watch other groups and we make sure we don't fall too far behind. If we do, we ask the TA for help or cut a few corners to get the product. (Undergrad 5)

Example 2 (O4-N6):

> No, they (labs) are really straightforward, we get a sheet about what we are doing, and he (instructor) goes over it and we copy the data we get onto the sheet so it's not too hard. (High School 6)

Theme 4: Criteria for Evaluating Own Work. The questions relating to Theme 4 probed the criteria that the participants used to evaluate their own chemistry work. The post hoc comparisons indicate that there were significant differences in responses on all pairwise comparisons except between graduate students and research trainees and between undergraduate and high-school students respectively (see Table 4). All the scientists gave O1 responses (see Table 5). These participants articulated multiple criteria for evaluating their own work. The criteria mentioned were (a) pragmatic factors such as the judgment of peers (acceptance in peer reviewed professional journals, grants, awards), the impact and fruitfulness of the work, and its relevance to society; (b) design criteria such as elegance, efficiency, and yield; and (c) cognitive factors such as the novelty, difficulty, and completeness of the work (see Table 2). Two excerpts illustrating these responses follow:

Example 1 (O1-N1):

> Well I guess I have I two driving forces . . . to make an environmental catalyst that is useful, cheap, and efficient . . . to develop new tools to look at surface chemistry which are kind of longer range projects, maybe a little bit more risky . . . I often pay attention as to whether . . . somebody really took the time to make a lot of careful measurements . . . as I progress I am finding more satisfaction in getting challenging experiments to work . . . it takes a lot of hard work to make something completely new beat other techniques that are more established . . . There are also certainly external markers . . . grant money or publications or recognition and invitations and all that kind of stuff . . . in [participant's area of research] . . . there are some very standard things you try and improve, sensitivity and resolution . . . (Scientist 9)

O2 responses focused exclusively on cognitive criteria such as the novelty, difficulty, or completeness of the work. These responses were by 68% of graduate students.

Example (O2-N2):

> . . . [It is good work if it is] novel chemistry or if you can use all the chemistry that you have, use known knowledge and apply [it] to new molecules, and also if you can finish something that people have not done before. (Grad 17)

O3 responses had a low-level task focus. As can be seen from Table 5, 90% of research trainees gave O3 responses. There were two variants of O3 responses (see Table 2). The participants who received a nominal code of N3 said that they evaluated their work based on whether or not they could complete their assigned laboratory tasks correctly and in a timely manner. The participants who received a nominal code of N4 said that their main criterion for evaluating their own laboratory work was whether their research supervisor or laboratory director was satisfied with their work. Specific examples follow:

Example 1 (O3-N3):

> How do I judge my work? . . . If things happen the way they should and I am able to get a decent product. I also think about how quickly or how many times it takes me to complete something . . . (Research Trainee 13)

Example 2 (O3-N4):

> When I get the results that he [laboratory director] wants, I consider that successful . . . (Research Trainee 8)

Ninety-four percent of the undergraduate and all of the high-school students gave O4 responses. These responses did not make any reference to empirical results or specific aspects of laboratory work. O4 responses included three variations represented by nominal codes N5, N6, and N7, respectively (see Table 2). The most common variant, N5, was given by all of the high-school students and 82% of the undergraduate chemistry students. These students essentially said that they evaluated themselves based on their ability to complete class assignments, course grades, and/or feedback from their teacher. A few undergraduates (12%) gave N6 responses. These students said that they judged their work based on whether or not it was of personal interest. Only one participant, a graduate student, gave an N7 response. She said that she did not like any of her work in science and found it uninteresting. A specific example of an O4 responses follows:

Example (O4-N5):

> The biggest thing I focus on is what Mr. [name deleted] says about how I am doing and my grades on homework and labs. (High School 18)

Theme 5: What Is Science? Questions for Theme 5 probed the participants' concepts of science in general. The participants were asked, "What is science?" Follow-up questions asked what differentiated science from nonscientific disciplines. Post hoc comparisons indicate that scientists differed significantly from every other group. There were no other significant differences (see Table 4). O1 responses were provided by 46% of scientists. These responses were multifaceted. The scientists started out by defining science as an attempt to understand the natural world. They said that what distinguished science from nonscientific endeavors was that science demanded objectivity based on standard procedures for gathering evidence. Further, they mentioned specific norms and standards for evaluating scientific claims including methodological norms (e.g., experimental manipulation, replication) and epistemic norms (e.g., precision and accuracy, replicability etc.). An example follows:

Example (O1-N1):

> Science is organized curiosity . . . I think the difference is whether or not you can . . . make measurements that other people can verify and in organic chemistry it is very seldom that anybody will ever do it [replicate others' work] to the extent of going

> the whole route but they will take little bits and pieces that they like, use it on their molecules and so that everything eventually gets tested in sort of orthogonal ways, so you don't usually get away with an outright lie. . . . Noticing unusual phenomena and trying to explain them . . . (Scientist 4)

A large number of participants from all five groups gave O2 responses (see Table 5). O2 responses were given by 54% of scientists, 77% of graduate students, 75% of research trainees, 47% of undergraduates, and 72% of high-school students. O2 responses described science as an attempt to understand the natural world. There were three variants of O2 responses at the nominal level (see N2, N3, and N4 in Table 2). N2 responses indicated that consensus on scientific knowledge was based on objective evidence or proof. However, these responses were not specific with regard to methodological and epistemic features that formed the basis of objective consensus in science. N3 responses focused on the cognitive aspects of scientific knowledge such as explanation, systematic analysis, etc. N4 was a unique response pattern, given by just one participant, a graduate student, revealing a relativistic stance with regard to scientific knowledge. Specific examples of O2 responses follow:

Example 1 (O2-N2):

> Science is a way to understand how things work . . . through observation and scrutiny of phenomenon in an objective, documented way . . . [in nonscience] There is not the same rigor or documentation of what you observe or it doesn't focus on how things work in the world." (Grad 21)

Example 2 (O2-N3):

> [Science is] logical thinking. . . It's looking at things in an orderly way, making sure what you are doing is systematic . . . (Grad 8)

Example 3 (O2-N4):

> . . . I think as scientists we think we have a lot more answers than we actually have. And we think that just because the numbers say something that there's truth associated with that. Where you know that if you've been in science long enough you can make the data say just about anything you want it to . . . Um, you know, whether you can get an answer besides or not is debatable in a lot of areas. (Grad 14)

O3 responses did not refer to the empirical content or methods of science. Instead, these responses either focused on the utilitarian aspects of science or defined science in terms of content (see N5 and N6 in Table 2). Twenty-three percent of graduate students, 25% of research trainees, 53% of undergraduates, and 26% of high-school students gave O3 responses.

Example 1 (O3-N5):

> I think science is a collection of ideas . . . that help us make sense of the world around us . . . Other areas, outside science, are not concerned with finding application for new things or developing more effective ways of keeping you healthy . . . (Research Trainee 12)

Example 2 (O3-N6):

> To me science is basically learning about how things work in nature, whether it is metabolism of a plant or that of a human or in general, not only that, as far as the

> anatomy or the physiology, it is just everything, you can look at the crux of the earth and try to figure out how a volcano erupts and what causes it to do that . . . (Research Trainee 3)

In summary, the descriptive data on group performance by theme showed that there were significant variations with expertise on each of the five themes examined in this study. The implications of our results for our key research questions will be discussed below.

DISCUSSION

This study was designed to answer three research questions about the development of epistemological knowledge in chemistry students. The implications of our results in relation to each of these research questions are discussed below.

Expertise and Epistemic Beliefs

The first research question we posed was: Do the participants' epistemic beliefs vary as a function of chemistry expertise? The results above clearly demonstrate that there was significant variation in epistemic beliefs as a function of chemistry expertise among the participants in this study. The results for Themes 2 (*Choice of Problems and Methods*), 3 (*Models for Handling Empirical Anomalies*), and 4 (*Criteria for Evaluating Own Work*) in particular, showed that those participants who were engaged in authentic chemistry research (scientists, graduate students, and undergraduate research trainees) conceptualized and evaluated their scientific work very differently from science students who lacked authentic research experiences (undergraduate and high-school students).

Most undergraduate and high-school students enrolled in chemistry laboratory courses did not articulate any criteria for knowledge evaluation that could be called "epistemic." Our findings about the impoverished views of science held by high-school and undergraduate students, whose main experience of science comes from conventional course and laboratory work, are consistent with previous research on students' epistemologies of science (Aikenhead, 1973; Carey & Smith, 1993; Ryder, Leach, & Driver, 1999). It is also consistent with recent work done by one of the authors that demonstrated a fundamental difference in the way chemistry majors view what happens in the classroom laboratory versus in the research laboratory (Del Carlo & Bodner, 2004).

We found important qualitative differences between scientists, graduate students, and undergraduate research trainees, indicating that differences in disciplinary expertise, especially in the duration and nature of participation in authentic research, had a significant impact on how the participants conceptualized science and scientific research. The scientists in our study, all highly respected and productive research chemists, responded to our questions in ways that were more complex, elaborated, and nuanced than any of the other groups. Their responses to each question were substantially longer than those of other respondents as might be expected from excerpts cited above.

More importantly, the practicing scientists typically articulated a variety of epistemic norms and pragmatic heuristics for evaluating knowledge claims in their discipline and in their own research. The specific norms and heuristics articulated by scientists overlapped considerably with a variety of epistemic and pragmatic criteria described in post-Kuhnian scholarship in the history and philosophy of science (Curd & Cover, 1998; Laudan et al., 1986). In contrast, even graduate (doctoral) students, who are one step away from becoming independent researchers, articulated a much more limited range of criteria for knowledge evaluation, focusing primarily on empirical adequacy criteria and methodological standards.

Another interesting difference among the groups was that only scientists saw a productive or generative research role for empirical anomalies. The faculty in this study repeatedly told us that some of their most exciting discoveries came from trying to understand and account for anomalous results. They also said that they typically planned sets of complementary experiments to gain a robust and accurate sense of the phenomenon being studied. Such triangulation across multiple experiments, specifically designed to control for known factors, and to contrast alternate plausible interpretations, helped the scientists identify and distinguish between routine technical error and productive anomalies in the course of their empirical research.

In contrast, graduate students and undergraduate research trainees tended to view empirical anomalies negatively. While they recognized that empirical anomalies were routine in laboratory work, they viewed these as a sign of their own technical limitations. Despite their awareness that empirical anomalies were frequent in the course of research, both of these groups of research-active students seemed to react to anomalous data after it occurred, rather than anticipating and planning for it in their studies.

In much the same way that the faculty and research-active graduate and undergraduate students differed in their perception of the role of empirical anomalies in laboratory work, there was a fundamental difference between the perception of anomalies held by the graduate students/research-active undergraduates and the undergraduate and high-school students whose view of "science" comes exclusively from exposure to classroom laboratory experiences. The participants in this study who had no exposure to authentic research practices often said that they had few or no problems with empirical anomalies in laboratory work. Some of these students also said that they dealt with the rare anomaly they encountered by fudging data or cheating. Once again, these results are consistent with prior work by one of the authors on student's perceptions of the classroom laboratory experience (Del Carlo & Bodner, 2004). Our conjectures about sources of these differences in participants' epistemologies will be discussed in relation to the next two research questions.

One aspect worth noting is that the results of the current study contrast with those of past researchers who have suggested that scientists may have relatively unsophisticated understandings of the nature of science (see, for example, Bell & Lederman, 2003). We believe that the main reason for these differences is our methodological approach, which tries to identify enacted epistemologies by anchoring epistemic questions in the participants' own practice and areas of research expertise.

Discipline-Specific Values and Research Heuristics

The second research question posed in this study was: Are there discipline-specific values and heuristics that guide chemistry research? From the responses of the scientists, and to a lesser extent of the graduate students, it was evident that disciplinary expertise and training in chemistry research contextualized the participants' epistemologies in important ways. Research chemists consistently offered contextualized accounts of their own work with reference to the norms and standards of disciplinary practice in chemistry. They expressed an awareness of the public or intersubjective dimensions of science (Husserl, 1973). They articulated several factors that shaped both their own scientific work and the way in which the scientific community of research chemistry produces, evaluates, and revises knowledge. For example they referred to (a) shared methodological standards of the chemistry research community, (b) the accuracy and replicability of empirical evidence, (c) the fit of their discoveries in the broader scheme of well-founded chemistry knowledge, (d) the fruitfulness of their research in terms of leading to new venues for research and discovery not only within their own programs but for their peers as well, (e) the relevance and utility of their work with respect to important social problems and needs, and (f) the judgment of research peers as expressed by the acceptance of their work in prestigious peer-reviewed journals, by external funding, and by professional awards.

Philosophers of science with diverse perspectives about the nature of science have suggested that many of these criteria serve to distinguish science from nonscience and are shared across the specific disciplines and subdisciplines of science (Kitcher, 1993; Kuhn, 1977; Laudan, 1990; Longino, 1990). However as noted in the introduction, contemporary scholarship on the nature of science suggests the specific content and form of the judgments that flow from these general criteria are shaped by discipline-specific factors or what Kuhn (1977) has referred to as the disciplinary matrix.

The responses of the practicing scientists clearly illustrate how both epistemic and pragmatic criteria for evaluating chemistry research are expressed and instantiated in terms of discipline-specific content. Indeed, the research chemists tended to spontaneously flesh out and elaborate their responses with specific examples from their own discipline, even when responding to the general questions about the nature of science in Theme 5. One scientist (Scientist 4), for example, described the role of empirical accuracy and replication by using an example of molecular synthesis when trying to define the nature of science (see example O1–Theme 5 under results).

While the graduate students did show some understanding of discipline-specific heuristics for evaluating their laboratory work, they typically focused on issues of empirical adequacy. They did not show any explicit awareness of how the public or intersubjective dimensions of science might shape epistemic or pragmatic criteria for knowledge evaluation. The other three groups of students did not demonstrate any clear understanding of these factors either.

The research chemists in our study also typically framed their scientific work in terms of a design or engineering language, rather than one of theory building and theory testing. For example, they talked about trying to build or synthesize molecules because chemistry theory predicted that such molecules could be built.

The scientists in this study articulated several discipline-specific design heuristics for evaluating their own scientific work. They talked about the efficiency and elegance of chemical synthesis, for example, based on such things as yield (how much of a target substance was actually obtained relative to the maximum amount possible on the basis of conservation of mass), the use of low-energy reactions, the difficulty of isolating intermediates in a reaction sequence, and the number of intermediary steps in the synthesis.

The ways which research chemists in this study thought about their scientific work did not correspond to certain "nature of science" concepts that prior science education research has focused on, such as the distinction between "theories" and "laws," or the nature of theory (see, for instance, Carey & Smith, 1993; Lederman et al., 2002; Smith et al., 2000). The research chemists did not appear to regard their own research as in any way "testing" the foundational theories of chemistry. Indeed, they typically reported that, even when faced with persistent empirical problems that required them to shelve a line of inquiry, they did not question the underlying body of chemistry theory. For example, no matter how many times and in how many different ways the chemists in this study tried to build a molecule and failed, they never questioned the body of chemistry theory which predicted that the molecule could be built. This is consistent with Kuhn's (1962) characterization of normal science.

We draw attention to these findings to raise larger questions about the goals and content of education about the nature of science. Some of the pedagogical objectives associated with the nature of science education have been (a) to increase understanding of what science is or to make people aware of the nature of science, (b) to increase scientific literacy or the ability of people to make informed decisions on issues involving science and technology, and

(c) to increase knowledge of the processes of scientific inquiry or how science is done (Bell et al., 2003). Our research suggests that unless a consideration of epistemic issues is closely linked to specific traditions of disciplinary knowledge and inquiry practices, students are unlikely to acquire sophisticated enacted epistemologies. This is important if one of the hopes of science education is that understanding the nature of science will help people reason better in their daily lives (Bell & Lederman, 2003; Lederman and O'Malley, 1990). Merely knowing in general that scientific knowledge is tentative rather than absolute or that scientific theories are supported by evidence for example, does not help a nonexpert evaluate the soundness of any particular knowledge claim in science.

Consider the question of how scientists know what a species is discussed earlier which has figured in some of the more recent attempts to assess epistemologies of science. Some recent experiments have shown that whether or not two groups of animals meet the criteria for being separate species is not fixed but depends on their environment. So for instance, two groups of cichlids that inhabit Lake Victoria (*Pundamila pundamilia* and *P. nyererei*) meet the standard definition for being separate species because although they occupy the same habitat and are "available" to each other as potential mates, they do not mate and are reproductively isolated. In fact, biologists regard these two varieties as separate cichlid species. However, when the environmental conditions are experimentally manipulated by placing both species in an artificial monochromatic light, the females mate across groups and produce reproductively viable offspring (Seehausen & van Alphen, 1998). One's evaluation of the "textbook" definition of species would clearly be influenced by access to such domain-specific knowledge.

We acknowledge that a broad understanding of the nature of science and scientific inquiry must include an understanding of how specific theories have changed in the history of the modern sciences and how they might change in the future. However, contemporary researchers in the history and philosophy of science note that theory-driven models do not necessarily account for the nature of scientific work in many areas of science such as biomedical research (Thagard, 2003, 2004). Indeed, there is agreement among many contemporary philosophers of science that the revision of foundational theories in any area of science is a rare event and much of routine scientific practice or inquiry is not really about testing foundational theories (Kuhn, 1962; Laudan et al., 1986; Thagard, 2004). Therefore, we suggest that one important goal of science education should be to help students understand the epistemic norms and values that govern current disciplinary practice within different areas of science.

We suggest that the current characterizations of the nature of science in science education may underrepresent important discipline-specific or contextual aspects of science and, in turn, hinder students' understanding of the epistemic and pragmatic bases for rational consensus in many areas of scientific practice. As noted above, one of the stated goals of education about the nature of science is to enhance scientific literacy and the ability of citizens to make informed practical decisions about specific scientific claims (for example with regard to the role of common chemicals in causing cancers). However, it seems to us that such informed decision making must be based on some contextualized understanding of how knowledge is created and evaluated in different areas of science. This is also true with regard to a third goal of education about the nature of science, which is to help students understand the processes of scientific inquiry. If important aspects of scientific inquiry processes vary by discipline within the sciences, students need to understand such variations in order to understand disciplinary inquiry.

While there is a growing awareness among science educators that there is no single "scientific method" (Bell et al., 2003), it is not clear that current attempts to teach the nature of science address the problem of domain specificity within the sciences in ways which will enhance students' enacted epistemologies. We will elaborate more on this point as we discuss the implications of our findings for our last research question.

The Nature of Research Experiences and the Development of Epistemic Beliefs

The third question our research addressed was: How does research experience influence the participants' epistemic beliefs? Our results indicate that differences in chemistry research expertise contributed significantly to the participants' epistemic beliefs. Our participants' responses suggested that the structure and demands of laboratory tasks, as well as the ways in which these participants' laboratory roles and responsibilities are defined, shaped their perspectives of what it meant to do science.

Research in situated cognition suggests that learning and sense making is always embedded in and shaped by a culture of practice (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991). The high-school and undergraduate chemistry students typically described their laboratory work as simple and highly structured, with worksheets specifying exactly what was to be done, observed, and recorded. The students believed that they knew the desired empirical outcomes of their laboratory work in advance and were given credit, primarily for (re)producing these outcomes. Even in such highly structured environments, one might expect empirical anomalies to be frequent, given their ubiquity in authentic research. However, it was clear that for many high-school and undergraduate students, classroom laboratory experiences were specifically designed to distract students from attending to such errors and focusing them instead on the process of filling out laboratory reports. These students often said that their instructors prompted them and gave them hints which helped them fill out "correct" results in their laboratory notebooks.

The problems associated with traditional laboratory work in high school and college classrooms have been documented before (Hake, 1998; Hegarty-Hazel, 1990; Hodson, 1993; Royuk & Brooks, 2003; Tobin & Gallagher, 1987). In response to these problems, many science education researchers have suggested that practical or laboratory work in science classrooms should involve "authentic" inquiry experiences (Chinn & Malhotra, 2002; Krajcik et al., 1998; Ryder & Leach, 1999). More recently, science educators have tried to teach students about the nature of science through research apprenticeship experiences that engage students in "real-world" inquiry, often in partnership with research scientists. However, these attempts have at best met with partial success (Barab & Hay, 2001; Bell et al., 2003; Carey & Smith, 1993; Leach, Hind, & Ryder, 2003; Smith et al., 2000). While these efforts to provide students with authentic inquiry experiences are clearly a move in the right direction, our findings with regard to undergraduate research trainees and graduate students both confirm and help explain the challenges of enhancing epistemic development through research apprenticeships.

Mere participation or apprenticeship in a real program of research does not guarantee that students will have opportunities to reflect on many aspects of scientific inquiry. Students who were engaged in authentic research told us that they learned a great deal about their research craft informally by asking questions and talking to their research group members, and even on occasion, to members of other research groups who might share research space and instruments. However, our research also indicated that the nature and focus of such informal conversations varied, depending on the research role of the participant and their position in the hierarchy of their research laboratory.

The undergraduate research trainees in our study could be viewed as initiates to the world of authentic chemistry research. This initiation clearly benefited their understanding of certain aspects of the nature of science. For example, the undergraduate research trainees shared the understanding of scientists and graduate students about the ubiquity of routine laboratory error. Unlike the undergraduate and high-school students enrolled in laboratory courses, the research trainees also began to develop rudimentary strategies for dealing with such error, typically in the form of retracing their steps and starting over.

Undergraduate research trainees were keenly aware of their peripheral participation in the overall research program. One such trainee referred to himself as "a small cog" in the wheel of research. They typically did not choose their specific research project, and had only a partial view of the big research picture because they were assigned the most routine tasks such as cleaning laboratory equipment, "prepping" or preparing instruments, or running simple chemical processes for other more senior research group members such as graduate students or postdoctoral students.

Often the undergraduate research trainees learned the technical aspects of their research craft from these more senior students rather than the scientists directly. Even when the undergraduate research trainees did talk to their supervising scientists, the conversations typically focused on solving specific technical problems that were holding up the trainees' work. Because of such communication hierarchies in research groups, undergraduate research trainees began to appreciate the ubiquity of empirical error, but rarely demonstrated a mastery of domain-specific heuristics for identifying and dealing with error.

In contrast, graduate students had more agency and control over their specific doctoral projects, and were expected to trouble shoot independently when they encountered empirical anomalies. They showed considerable knowledge of domain-specific heuristics for evaluating their empirical work and identifying sources of routine laboratory error. However, funding for research and strategic decisions to pursue or drop specific lines of empirical inquiry were controlled by the research scientists with whom graduate students worked. Consequently, the graduate students typically showed little awareness of the pragmatic factors that went into the strategic direction and redirection of research programs by scientists. This might be one reason why graduate students, in contrast to scientists, showed little awareness of the productive possibilities of anomalies. While scientists, having strategic and financial control over their overall program of research, could choose to redirect their attention or to pursue entirely new research questions that arose from productive anomalies, such choices were typically not the graduate students' to make.

On the whole, our research indicates that immersion in authentic research experiences provides students with important opportunities to learn about the processes of scientific inquiry specific to their discipline. Gains in research expertise appear to come with sustained immersion and with more central participatory roles in authentic research.

Limitations and Directions for Future Research

Our findings suggest that an important aspect of epistemic development in chemistry is the acquisition of functional, discipline-specific epistemic and pragmatic heuristics used by practicing scientists in a given discipline that guide their ongoing research and help them evaluate the research of their peers. Additionally, we believe that sustained apprenticeship in real-world research is important to chemistry students' epistemic development. However, further research is needed to understand the specific processes of epistemic development and how research experiences might be enhanced to provide junior apprentices more insight into the nature of inquiry in their specific disciplines. The current data were drawn from interviews with the participants. In future research, observations of the participants' laboratory work, formal research group meetings, and informal conversations among research group members would provide important data on how students' roles and research tasks shape their epistemic beliefs.

Additionally, our findings about the contextual nature of epistemic development must be qualified, as we did not observe participants across different scientific disciplines. It would be interesting to examine whether the current patterns of results could be replicated with research in other scientific disciplines.

Conclusions and Implications for Science Education

Our research indicates that research chemists use discipline-specific, functional, epistemic and pragmatic heuristics to evaluate chemistry knowledge. The ability of chemistry students to understand and use such epistemic and pragmatic heuristics depends on the nature and duration of their experience with authentic chemistry research. The nature of students' research participation, as defined by their research roles, influences how they interact with more expert group members and their opportunities to experience and solve research problems. In turn, such variations in research experience influence students' epistemic development. One implication of these findings is that apprenticeship experiences in authentic inquiry do not automatically guarantee epistemic development in students. Opportunities for epistemic development will most likely depend on the degree of research autonomy given to students, and on opportunities to engage with expert researchers in conversations and reflections on epistemic issues related to the program of research.

REFERENCES

Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. Science Education, 82, 417–436.

Abd-El-Khalick, F., & Lederman, N. G. (2000). The influence of history of science courses on students' views of nature of science. Journal of Research in Science Teaching, 37(10), 1057–1095.

Aikenhead, G. (1973). The measurement of high-school students' knowledge about science and scientists. Science Education, 57, 539–549.

Alters, B. J. (1997). Whose nature of science? Journal of Research in Science Teaching, 34(1), 39–55.

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

Barab, S. A., & Hay, K. (2001). Doing science at the elbows of scientists: Issues related to the scientist apprentice camp. Journal of Research in Science Teaching, 38(1), 70–102.

Bell. R. L., Blair, L. M., Crawford, B. A., & Lederman, N. G. (2003). Just do it? Impact of a science apprenticeship program on high-school students' understanding of the nature of science and scientific inquiry. Journal of Research in Science Teaching, 40(3), 487–509.

Bell, R. L., & Lederman, N. G. (2003). Understandings of the nature of science and decision making on science and technology based issues, Science Education, 87, 352–377.

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

Carey, S., Evans, R., Honda, M., Unger, C., & Jay, E. (1989). An experiment is when you try and see if it works: Middle school conception of science. International Journal of Science Education, 11, 514–529.

Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. Educational Psychologist, 28(3), 235–251.

Chi, M. T. H. (1997). Quantifying qualitative analysis of verbal protocol data. The Journal of the Learning Sciences, 6(3), 271–315.

Chinn, C. A., & Malhotra, B. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. Science Education, 86(2), 175–218.

Chinn, C. A., & Samarapungavan, A. (2005). Toward a broader conceptualization of epistemology in science education. Paper presented at the meeting of the Eighth International History, Philosophy, Sociology, and Science Conference. University of Leeds, England.

Conant, J. B. (1957). The overthrow of the phlogiston theory. The chemical revolution of 1775–1789. In J. B. Conant & L. K. Nash (Eds.), Harvard case histories in experimental science (pp. 67–115). Cambridge, MA: Harvard University Press.

Curd, M., & Cover, J. A. (Eds.). (1998). Philosophy of science: The central issues. New York: Wiley.

Del Carlo, D., & Bodner, G. M. (2004). Students' perceptions of academic dishonesty in the chemistry classroom laboratory. Journal of Research in Science Teaching, 41(1), 47–64.
Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J. Sternberg & J. Davidson (Eds.), Mechanisms of insight (pp. 365–395). Cambridge, MA: MIT Press.
Duschl, R. A. (1988). Abandoning the scientistic legacy of science education. Science Education, 72, 51–62.
Duschl, R. A. (2000). Making the nature of science explicit. In R. Millar, J. Leach, & J. Osborne (Eds.), Improving science education: The contribution of research (pp. 187–206). Philadelphia, PA: Open University Press.
Giere, R. N. (1988). Explaining science: A cognitive approach. Chicago, IL: University of Chicago Press.
Goldman, A. (1986). Epistemology and cognition. Cambridge, MA: Harvard University Press.
Goodman, N. (1983). Fact, fiction, and forecast. Cambridge, MA: Harvard University Press.
Hake, R. R. (1998). Interactive engagement versus traditional methods: A six thousand student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66, 64–74.
Hegarty-Hazel, E. (1990). The student laboratory and the science curriculum. London: Routledge.
Hodson, D. (1985). Philosophy of science and science education.Studies in Science Education, 12, 25–57.
Hodson, D. (1988). Towards a philosophically more valid science curriculum. Science Education, 72, 19–40.
Hodson, D. (1993). Re-thinking old ways: Towards a more critical approach to practical work in school science. Studies in Science Education, 22, 85–142.
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.
Hogan, K. (2000). Exploring a process view of students' knowledge about the nature of science. Science & Education, 84(1), 51–70.
Hollander, M., & Wolfe, D. A. (1999). Nonparametric statistical methods (2nd ed.). New York: Wiley.
Husserl, E. (1973). Experience and judgment (J. S. Churchill & K. Ameriks, Transl.). London: Routledge.
Kitcher, P. (1992). The naturalists' return. Philosophical Review, 101, 53–114.
Kitcher, P. (1993). The advancement of science: Science without legend, objectivity without illusions. Oxford, UK: Oxford University Press.
Krajcik, J. S., Blumenfeld, P., Marx, R. W., Bass, K. M., Fredericks, J., & Soloway, E. (1998). Middle school students' initial attempts at inquiry in project-based science classrooms. Journal of the Learning Sciences, 7, 313–350.
Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association, 47(260), 583–621.
Kuhn, T. S. (1962). The structure of scientific revolutions. Chicago, IL: University of Chicago Press.
Kuhn, T. S. (1977). The essential tension. Chicago, IL: University of Chicago Press.
Kuhn, T. S. (2000). The road since structure: Philosophical essays, 1970–1993 (J. Conant & J. Haugeland, Eds.). Chicago, IL: University of Chicago Press.
Laudan, L. (1984). Science and values. Berkeley, CA: University of California Press.
Laudan, L. (1990). Science and relativism: Some key controversies in the philosophy of science. Chicago, IL: Chicago University Press.
Laudan, L., Donovan, A., Laudan, R., Barker, P., Brown, H., Leptn, J., Thagard, P., & Wykstra, S. (1986). Scientific change: Philosophical models and historical research. Synthese, 69, 141–223.
Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge, UK: Cambridge University Press.
Lawson, A. E. (2000). The generality of hypothetico-deductive reasoning: Making scientific reasoning explicit. The American Biology Teacher, 62(7), 482–495.
Leach, J., Hind, R., & Ryder, J. (2003). Designing and evaluating short interventions about the epistemology of science in high school science classrooms. Science Education, 87, 831–848.
Leach, J., Millar, R., Ryder, J., & Séré, G. M. (2000). Epistemological understanding in science learning: Consistency of representation across contexts. Learning and Instruction, 10(6), 497–527.
Lederman, N. G. (1986). Students' and teachers' understanding of the nature of science: A reassessment. School Science and Mathematics, 86, 91–99.
Lederman, N. G., & O'Malley, M. (1990). Students' perceptions of tentativeness in science: Development, use, and sources of change. Science Education, 74, 225–239.
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.
Linn, M. C., & Songer, N. B. (1993). How do students make sense of science? Merrill-Palmer Quarterly, 39(1), 47–73.
Longino, H. (1990). Science as social knowledge: Values and objectivity in scientific inquiry. Princeton, NJ: Princeton University Press.

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

Nersessian, N. (1984). Faraday to Einstein: Constructing meaning in scientific theories. Dordrecht, The Netherlands: Kluwer.

Patton, M. Q. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park, CA: Sage.

Perry, W. G. (1970). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rinehart & Wilson.

Popper, K. (1959). The logic of scientific discovery. London: Hutchinson.

Popper, K. (1972). Objective knowledge: An evolutionary approach. Oxford, UK: Oxford University Press.

Quine, W. V. (1970). Philosophy of logic. Englewood Cliffs, NJ: Prentice-Hall.

Royuk, B., & Brooks, D. W. (2003). Cookbook procedures in MBL physics exercises. Journal of Science Education and Technology, 12(3), 317–324.

Ryder, J., & Leach, J. (1999). University science students' experiences of investigative project work and their images of science. International Journal of Science Education, 21(9), 945–956.

Ryder, J., Leach, J., & Driver, R. (1999). Undergraduate science students' images of science. Journal of Research in Science Teaching, 36(2), 201–219.

Sahdra, B., & Thagard, P. (2003). Procedural knowledge in molecular biology. Philosophical Psychology, 16, 477–498.

Samarapungavan, A. (1992, April). Scientists' conceptions of science: A study of epistemic beliefs. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Samarapungavan, A. Westby. E. L., & Bodner, G. M. (1999, April). Scientist as engineers: Epistemic beliefs and values in the applied sciences. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Samarapungavan, A., & Wiers, R. (1997). Children's thoughts on the origin of species: A study of explanatory coherence. Cognitive Science, 21(2), 147–177.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. Journal of Educational Psychology, 82, 498–504.

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

Seehausen, O., & van Alphen J. J. M. (1998). The effect of male coloration on female choices in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). Behavioral Ecology and Sociobiology, 42, 1–8.

Smith, C. L., Maclin, D., Houghton, C., & Hennessy M. G. (2000). Sixth grade students' epistemologies of science: The impact of school science experiences on epistemological development. Cognition & Instruction, 18(3), 349–422.

Steele, R. G. D. (1960). A rank sum test for comparing all pairs of treatments. Technometrics, 2, 197–207.

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

Thagard, P. (1990). The conceptual structure of the chemical revolution. Philosophy of Science, 57, 183–209.

Thagard, P. (2003). Pathways to biomedical discovery. Philosophy of Science, 70, 235–254.

Thagard, P. (2004). Rationality and science. In A. Mele & P. Rawlings (Eds.), Handbook of rationality (pp. 363–379). Oxford, UK: Oxford University Press.

Tobin, K., & Gallagher, J. J. (1987). What happens in high school science classrooms? Journal of Curriculum Studies, 19, 549–560.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. Cognitive Psychology, 24, 535–585.

Westby, E. L. (2003). The influence of disciplinary expertise on epistemic beliefs in chemistry. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Westby, E., Samarapungavan, A., & Bodner, G. M. (2000, April). Strategic decision making about the direction of research programs: A comparison of decision making by chemistry graduate students and research faculty in applied chemistry. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.