Research Article

Assessing Students’ Understanding of Inquiry: What Do Prospective Science Teachers Notice?

Vicente Talanquer, Debra Tomanek, and Ingrid Novodvorsky

College of Science Teacher Preparation Program, University of Arizona, P.O. Box 210106, Tucson, Arizona 85721-0106

Received 25 July 2011; Accepted 16 December 2012

Abstract: The theoretical construct of teacher noticing has allowed mathematics teacher educators to examine teacher thinking and practice by looking at the range of activities that teachers notice in the classroom. Guided by this approach to the study of teacher thinking, the central goal of this exploratory study was to identify what prospective science teachers notice when evaluating evidence of student understanding in another teacher’s inquiry-based unit. Our results are based on the qualitative analysis of 43 prospective teachers’ evaluations of assessment evidence presented to them in the form of a video case and associated written artifacts. Analysis of our data revealed two major categories of elements, Task-General and Task-Specific, noticed by our study participants. Task-General elements included attention to learning objectives, independent student work, and presentation issues and they often served to guide or qualify the specific inquiry skills that were evaluated. Task-Specific elements included the noticing of students’ abilities to perform different components of an investigation. In general, study participants paid attention to important general and specific aspects of student work in the context of inquiry. However, they showed preferential attention to those process skills associated with designing an investigation versus those practices related to the analysis of data and generation of conclusions. Additionally, their interpretations of assessment outcomes were largely focused on the demonstration of general science process skills; much less attention was paid to the analysis of the epistemological validity or scientific plausibility of students’ ideas. Our results provide insights into the design of meaningful learning experiences for prospective teachers that elicit, challenge, and enrich their conceptions of student understanding in the context of inquiry. © 2013 Wiley Periodicals, Inc. J Res Sci Teach 50: 189–208, 2013

Keywords: assessment; inquiry; prospective science teachers; teacher thinking; teacher noticing

Reforms in science education over the past 20 years have emphasized the importance of engaging students in inquiry by creating opportunities for them to generate and evaluate scientific explanations of the natural world, and to participate in scientific practices and discourse (AAAS, 1993; NRC, 1996, 2007; Osborne & Dillon, 2008). However, educational research has elicited the many challenges that practicing and prospective science teachers face to successfully design and implement effective inquiry-based learning experiences. These studies have explored the influence of personal conceptions and beliefs about the nature of science (Roehrig & Luft, 2004), the purpose of education (Lotter, Harwood,
& Bonner, 2007), the nature of effective teaching practices (Crawford, 2007), and the abilities and motivation of students (Wallace & Kang, 2004) on the type and amount of inquiry-based instruction enacted in the classroom. However, we know little about how science teachers approach the assessment of inquiry and how they interpret evidence of student understanding in such contexts. The present exploratory study was designed to increase our knowledge in these areas.

Although classroom assessment has become a prime focus of educational researchers and policy makers in recent years (Black & Wiliam, 1998; NRC, 2000, 2001), research on teachers’ knowledge and reasoning with regard to assessment is limited. Most studies have focused on describing teachers’ knowledge and beliefs about assessment (Graham, 2005; Maclellan, 2004; Ruiz-Primo & Shavelson, 1996; Wang, Kao, & Lin, 2010), cataloguing the types of assessment tools most commonly used by teachers (Campbell & Evans, 2001), or analyzing teachers’ general assessment practices (Sato, Wei, & Darling-Hammond, 2008). Little is known about teachers’ assessment reasoning, that is, the reasoning that teachers follow when selecting, implementing, and interpreting the results of assessment tasks designed to target specific understandings in a discipline. Yet, this knowledge could facilitate the design of meaningful learning experiences for both practicing and prospective teachers that would elicit, challenge, and enrich their conceptions about assessment and help them develop an evidence-based approach to assessing student learning and understanding.

We believe that prospective teachers’ assessment reasoning is an especially critical, yet underdeveloped, area of their professional development (Shepard, 2000). Given the scarcity of research in this area, our efforts in recent years have focused on better understanding the nature and development of secondary school science teachers’ thinking about assessment. We have, for example, explored the main factors that influence prospective and experienced science teachers’ reasoning as they select tasks to formatively assess their students’ understanding of scientific concepts (Tomanek, Talanquer, & Novodvorsky, 2008). This analysis revealed that teacher thinking in this area seems to be largely influenced by characteristics of the assessment tasks related to cognitive demand and perceived efficacy (e.g., focus on targeted concepts), as well as by characteristics of the students, such as perceived ability to complete the assessment task. In this article, we shift our focus of attention from the analysis of how science teachers go about selecting assessment tasks to investigating how they approach the analysis of evidence of student understanding. In particular, we investigated what prospective science teachers noticed as they evaluated student work in a multi-day, inquiry-based science unit. The theoretical construct of teacher noticing has allowed mathematics teacher educators to examine teacher thinking and practice by looking at the range of activities that teachers do and do not notice in the classroom (Sherin, Jacobs, & Philipp, 2011). This construct has also been useful in guiding the analysis of how teachers evaluate student work (Goldsmith & Seago, 2011; Kazemi & Franke, 2004).

The core findings of the present study are based on the qualitative analysis of over 40 prospective science teachers’ evaluations of assessment evidence presented to them in the form of a video case and associated written artifacts. Our results elicit critical aspects of student work that influenced prospective teachers’ interpretations of student understanding, and reveal patterns in noticing that may be indicative of deep beliefs about inquiry or of how prospective teachers frame the task of evaluating student understanding in the context of inquiry. From this perspective, our findings provide insights into how to support prospective science teachers’ learning and help them become better assessors of inquiry-based learning experiences.

Journal of Research in Science Teaching
Assessment Reasoning and Practices in the Context of Inquiry

Existing research findings suggest that teachers’ assessment reasoning is often constrained by their prior conceptions of student learning and assessment (Brookhart, 2004; Otero, 2006), beliefs about fairness (Yung, 2001), pressures of accountability (Brickhouse & Bodner, 1992; Duschl & Wright, 1989), and concerns for maintaining supportive relationships with students (Fieman-Nemser & Floden, 1986). In fact, many of the barriers to teachers’ implementation of assessments compatible with teaching for understanding appear to be psychological and social, rather than technical or psychometric (Dwyer, 1998). Some researchers have suggested that teachers’ assessment practices mirror their limited beliefs about assessment as a tool for measuring achievement rather than assessment as a means of revealing student understanding and improving student learning (Gioka, 2009; Radnor, 1994; Shepard, 2000).

A few researchers have investigated science teachers’ views about assessing students’ investigative skills or their abilities to engage in inquiry. Findings from these studies indicate that secondary school teachers tend to view such skills in terms of consecutive steps, laying most of their assessment emphasis on two areas: the formulation of the research question and the drawing of conclusions (Stokking, Schaaf, Jaspers, & Erkens, 2004). Analysis of initial conceptions of assessment among prospective elementary school teachers has revealed that only a small fraction of the teachers seem to value or consider assessment of the process of inquiry in their regular practice (Wang et al., 2010). Of these teachers, most tend to focus their attention on the assessment of procedural abilities needed to carry out investigations, such as observation, measurement, and data recording skills.

Duschl (2003) has suggested that the assessment of inquiry should focus on three integrated domains: (1) Epistemic—This domain includes the abilities involved in the processes of science (i.e., science process skills such as observing, measuring), but also the knowledge structures and the criteria used in science to make critical judgments about the products of inquiry; (2) Conceptual—This domain includes the conceptual frameworks that students are expected to develop and the alternative conceptions that they bring with them to the classroom; (3) Social—This domain refers to the representation and communication frameworks students use while engaging in inquiry. The social domain includes the processes that shape how knowledge is represented, communicated, argued, and debated. Research findings by Ruiz-Primo and Furtak (2007) suggest that science teachers may emphasize some aspects of the epistemic domain (mainly process skills) over the conceptual domain in their informal formative assessment practices in the context of scientific inquiry. In particular, a large proportion of the eliciting questions asked by middle school teachers in this latter study were devoted to the application of known procedures and to the description of observations rather than to the formulation of explanations and the evaluation of the quality of evidence.

Recently, Tang, Coffey, Elby, and Levin (2010) have suggested that a focus on the scientific method as a discrete set of independent steps may draw teachers’ attention away from instances of productive inquiry in their classrooms. At both the individual and group level, focused attention on the correctness of students’ use of variables or formulation of hypotheses caused teachers to overlook productive aspects of students’ reasoning, such as attending to confounding causal factors, generating mechanistic explanations, and attending to data reliability. Other authors (Russ, Coffey, Hammer, & Hutchison, 2009) have also pointed out that science teachers tend to evaluate students’ ideas in the context of inquiry based on how well students’ answers and procedures compare to the canon as represented by the curriculum, undervaluing and undermining other productive ways of reasoning. Research suggests that
novice science teachers can begin to attend to student thinking early in their teacher careers, but they frequently work in professional environments that focus their attention on other issues (Levin, Hammer, & Coffey, 2009).

Teacher Noticing

In recent years, a significant body of research in mathematics teacher education has focused on the exploration and analysis of what teachers notice or attend to in the classroom (Sherin, Jacobs, et al., 2011). Research guided by the theoretical construct of teacher noticing has elicited differences in what and how novice and experienced mathematics teachers monitor, interpret, and recall when viewing video of instruction (Sherin, 2007). Results from these types of studies suggest that teachers’ attention tends to be more focused on issues of classroom organization and instruction than on student thinking about the content under discussion (Sherin & Han, 2004). However, these studies have also shown that noticing is: (a) selective, as it often involves attending to some events and disattending to others; (b) multidimensional, as teachers often have diverse objects of attention; and (c) instrumental, as what is attended to is usually what requires action by the teacher. In general, noticing patterns seem to be strongly influenced by a teacher’s experiences, prior knowledge, expectations, and pedagogical commitments (Erickson, 2011).

Researchers have considered different aspects of teacher thinking and practice in their definition of noticing (Sherin, Russ, & Colesstock, 2011). For example, some of them conceive noticing as involving only what the teachers attend to when they view a classroom lesson (Star & Strickland, 2008). Other researchers are interested not only in what catches a teacher’s attention but also in teachers’ interpretations of what is noticed (van Es & Sherin, 2002). In some cases, the definition of noticing may also include teachers’ plans to respond to what is noticed. The analysis of what teachers’ notice may thus vary from one type of study to another. However, researchers in this field tend to reflect on issues such as the object of noticing (e.g., teacher action, student thinking, content issue), the noticing stance (e.g., descriptive, evaluative, interpretative), the specificity of noticing (e.g., specific student, whole class), and the noticing focus (e.g., specific concept, general topic) (van Es & Sherin, 2008).

Much of the research on teacher noticing has been based on the analysis of interviews in which teachers are asked to reflect on what they noticed after teaching a class (Ainley & Luntley, 2007) or after watching a video of instruction (Colesstock & Sherin, 2009). This latter methodology has the advantage of not relying on recalled events and self-reports. In these different studies, researchers have frequently focused their analysis on the identification of patterns of noticing, typically by documenting the percentages of teachers or teachers comments that fall into different categories (Sherin & van Es, 2009; Star & Strickland, 2008). This approach helps to elicit typical behaviors for a certain group of teachers (e.g., novice versus experienced teachers), but does not capture variations in those patterns with time and varying contexts. Recently, some authors have used portable video cameras that allow teachers to capture noticing from their own perspective in a moment-to-moment basis (Sherin, Russ, et al., 2011).

Although most of the research on teacher noticing has been done in the context of mathematics teacher education, recent studies in science education have begun to highlight the importance of carefully analyzing what science teachers pay attention to, particularly in reference to student thinking (Coffey, Hammer, Levin, & Grant, 2011; Hammer & van Zee, 2006; Roth et al., 2011; Russ et al., 2009). These types of studies emphasize the importance of analyzing what prospective and in-service science teachers attend to in order to create

Journal of Research in Science Teaching
learning opportunities that better support teachers’ ability to genuinely engage with students’ ideas in the classroom.

Goals and Methodology

Study Goals

The central goal of this exploratory study was to identify patterns of noticing in our prospective science teachers’ evaluation of student work in an inquiry-based science unit. In particular, our data analyses were guided by the following research question:

- What do prospective science teachers notice in their evaluation of student understanding in the contexts of an inquiry-based unit of instruction?

We were particularly interested in identifying those elements of student work most commonly used by our prospective teachers in making claims about the quality of the understanding developed by students engaged in inquiry.

Setting and Participants

This study was conducted at a public university in the southwestern United States. Our research participants included prospective science teachers \((N = 43; 33 \text{ female, } 10 \text{ male})\) enrolled in an undergraduate secondary science teacher preparation program. Students in this program are science or science education majors who complete their teacher training in the College of Science (Talanquer, Novodvorsky, Slater, & Tomanek, 2003). Science education courses in our program include an introductory course on general issues on teaching science in secondary classrooms and a course on how people learn. Our prospective teachers also take one advanced course on planning, instruction, and assessment issues in the sciences, a course on managing the learning environment, and a subject-specific (biology, chemistry, or physics) teaching methods course. All of these courses put a strong emphasis on the analysis and reflection of teaching and assessment practices that engage students in inquiry and focus on student understanding of central ideas in the sciences. As part of these different courses, prospective teachers spend close to 200 hours working in secondary school classrooms under the supervision of experienced mentor teachers. During these internships, prospective teachers have several opportunities to observe and engage in planning, implementing, and assessing science lessons with different levels of inquiry. All of the participants in our study had satisfactorily completed all of these science education courses and were involved in a one-semester student teaching experience in local secondary school classrooms.

Research Instruments and Data Collection

Our results are based on the analysis of a portion of the data collected using an assessment instrument designed to assess our prospective teachers’ ability to reflect upon and evaluate another science teacher’s planning, instruction, and assessment decisions. All of our prospective teachers are asked to complete this assessment 7 or 8 weeks into their one-semester long student teaching experience as part of a global evaluation of their teaching abilities. The assessment has a video case format and is tailored to the prospective teachers’ areas of science preparation. The complete packet of information provided in the assessment is therefore different for the prospective teachers depending upon the area of specialty in which each is prepared (different assessment task packets for biology, chemistry, and physics). To complete the assessment, prospective teachers are given access to a video lesson and to a set of
associated artifacts (e.g., teacher’s plans, samples of student work) that they can use as evidence to support the claims they made in their analysis. For example, the video lesson and accompanying artifacts given to our future biology teachers corresponds to a unit of instruction in which students are engaged in an inquiry-based learning experience about cricket behavior.

The present study used data collected from all of the prospective biology teachers who have completed the last version of our “Decision-Making Assessment” since it was implemented 6 years ago (N = 43). Although the pedagogical approaches and content emphases have changed slightly in the science teacher preparation program curriculum, a consistent focus on inquiry-based lessons has remained throughout the years in which the study subjects participated as students in the program. We analyzed only the assessment task responses from the biology teachers, and not the chemistry or physics teachers, to limit the variation in responses due to the differences in the subject matter and purposes contained in the three versions of the video lesson/artifacts assessment instrument. The biology teachers were also the largest group of the three subject area data sets. Given our research interests, we only considered prospective teachers’ responses to the last of four tasks in this assignment which focuses on the analysis of student understanding. The prompt for this specific task is shown below:

**Task 4. Analyzing Student Understanding:** This packet includes several forms of evidence that can be used to infer student understanding (e.g., student discussion in videos, transcripts of student discussions, samples of written student work). Select two forms of evidence to analyze and evaluate the extent to which student understanding is demonstrated in each form.

We believed that prospective teachers’ written responses to this task would allow us to explore what our prospective teachers notice in evaluating student understanding in the context of inquiry. As commonly done in research studies on teacher noticing, the task was based on the analysis of a video of instruction and a sample of classroom artifacts (Sherin, Jacobs, et al., 2011). In particular, the “Decision Making Assessment” packet for our prospective biology teachers includes the following resources:

- The video “Teaching high school science: Investigating crickets” from the Annenberg/CPB Collection (2000). This video depicts ninth-grade biology students in Ms. McClellan’s class as they design and conduct experiments about crickets in collaborative pairs. The video shows days 2, 3, and 4 in the 7-day unit and it depicts Ms. McClellan’s introduction to the unit, some discussions between students while engaged in designing and conducting experiments, as well as several discussions between the teacher and student teams as the class is engaged in the inquiry activities.
- A one-page transcript of a brief conversation between the teacher and two students on Day 2 of the video; the conversation focuses on the identification of relevant experimental variables in the study.
- A two-page lab journal entry for this investigation from one student in the class. This entry includes descriptions of the purpose of the study, the variables to test, the initial and final hypotheses, the materials used, the experimental design, and a data table with results.
- A final poster from a group of students. The poster includes sections describing the purpose of the study, the initial hypothesis, the experimental procedure, a data table with results, the discussion of results, and the conclusions.
- A one-page entry from the teacher’s journal regarding her post-class teaching reflections. In these reflections, the teacher expresses concerns about her students’ ability
to link the different parts of their investigations and also about the focus of their observations.
- A 7-day plan for the unit entitled “Nature investigation: Cricket behavior.”
- A timeline of units comprising the 9th grade biology class that Ms. McClellan teaches.

The first four items in this list provide a direct avenue to student work and thinking during the unit of instruction, while the last three artifacts can be used to infer the teacher’s goals, expectations, and personal evaluations of student work. The packet for this assignment also includes a description of the teaching context in which the 7-day unit was implemented and the set of four tasks that prospective teachers are expected to complete based on the evidence provided. All of the above resources, except for the video, are included for reference in the Supporting Information to this article. We encourage readers to review these materials in order to develop a better sense of the breadth and depth of the evidence available to prospective teachers in the completion of this assignment.

Data Analysis

Prospective teachers’ written responses to task 4 of our Decision Making Assessment were analyzed using a constant comparison method looking for emerging patterns (Miles & Huberman, 1994). Each of the authors separately read through each of the 43 written answers and identified specific codes to describe the different elements noticed and used by prospective teachers to make claims about student understanding. The individual descriptive codes that emerged from this analysis, as well as their application to each of the prospective teachers’ responses, were then reviewed and discussed by the authors, looking for consensus on a set of descriptors that best characterized each of the answers. Examples of such descriptive codes included “clarity,” “identify proper question,” and “controlling variables.” As a second part of the analysis, descriptive codes were further discussed by the investigators and consensus was reached about grouping these codes into general categories describing different dimensions of noticing. This work was guided by prior research in the area of teacher noticing (Goldsmith & Seago, 2011; van Es & Sherin, 2008). Following is a description of the major coding categories that emerged from the analysis.

The aspects or elements of student work noticed by prospective teachers during their evaluations were classified into two major categories: Task-General and Task-Specific. Task-General included elements that could be noticed in the evaluation of any type of school task, while Task-Specific referred to elements that were characteristic of the task at hand (inquiry-based unit). The Task-General elements were in turn subdivided into three subcategories: Learning Objectives, Independent Work, and Presentation. Learning Objectives refers to paying explicit attention to the stated goals of the unit in making judgments about student understanding (e.g., “This indicates to me that these two students understood enough to fulfill the class objectives” or “She has written the criteria that is required of her”). Independent Work refers to noticing the extent to which outcomes are the result of individual effort (e.g., “The table was created solely by the students” or “It was then that, unguided, the students were not able to correctly present their conclusion”). Presentation refers to attending to the general format, organization, completeness, and clarity of student work (e.g., “The poster had all of the information of the experiment” or “All parts of the lab set up and data collection are labeled”).

Given the nature of the Task-Specific elements noticed by the prospective teachers in our sample, we opted for subdividing these elements into five groups: Asking Questions and
Formulating Hypotheses, Designing and Setting Up Experiments, Observing and Recording, Interpreting and Evaluating Evidence, and Aligning Components, based on the type of science practice to which they belonged (NRC, 2007). Asking Questions and Formulating Hypotheses refers to noticing issues related to students’ abilities to formulate a research question, understand and identify experimental variables, or build testable hypotheses (e.g., “The student really had a handle on what a variable is” or “They did not know how to write and if . . . then . . . hypothesis”). Designing and Setting Up Experiments refers to attending to students’ ability to control variables in their experiments or to students’ general ability to set up an experiment (e.g., “Understanding how to design an experiment with only one variable is apparent in their work” or “They run multiple trials and include a control environment”). Observing and Recording refers to noticing students’ abilities to collect and organize experimental data (e.g., “They did not include any impertinent observations” or “Data collection and display could use improvement”). Interpreting and Evaluating Evidence refers to paying attention to students’ ability to build explanations and generate adequate conclusions (e.g., “There are no deep explanations of why the results were the way they turned out” or “The students concluded something they weren’t testing for”). Finally, Aligning Components refers to noticing issues of coherence between different components of students’ investigations (e.g., “Their hypothesis was well aligned with the procedure and the data collection” or “Their procedure is a simple, yet suitable set up to test their variable”).

As part of our analysis, we kept track of the number of instances in which a given type of element was identified within the prospective teachers’ evaluations of each type of assessment evidence provided. Tables were created indicating the number and percent of prospective teachers who noticed elements in the different coding categories described in the previous paragraphs. Whether a prospective teacher noticed an element in one, two, or more types of assessment, the teacher was counted just once. This quantitative analysis allowed us to identify patterns in teachers’ noticing while engaged in the evaluation of student work. We also extracted qualitative data in the form of excerpts from prospective teachers’ responses to the Decision-Making task to better explain or support our assertions. Excerpts of responses from a given prospective teacher were assigned the same alphanumerical label (e.g., DM1).

Findings

Prospective teachers paid attention to two main categories of factors during their evaluation of student understanding in the context of an inquiry-based learning experience: (1) Task-General and (2) Task-Specific. Task-general factors served to guide or qualify prospective teachers’ evaluations (e.g., “this is evidence she has met learning objective 4”) of the task-specific features used to assess student understanding of inquiry (e.g., “ability to formulate a research question”). Thus, prospective teachers’ evaluations of student understanding were influenced by both general expectations of performance as well as specific judgments of inquiry skills. The remainder of this section presents a description of the different types of noticed elements and of their influence on prospective teachers’ assessment thinking.

Task-General Elements

Prospective teachers consistently noticed task-general elements belonging to three main categories: Learning Objectives, Independent Work, and Presentation. Over three quarters (33/43, 76.7%) of all of the study participants noticed at least one task-general element in their evaluation of student understanding. The characterization of these types of elements is important as they guided, and also constrained, prospective teachers’ assessment reasoning.
Learning Objectives. Close to a third of the study participants (13/43, 30.2%) explicitly referred to the learning objectives included in the “Decision Making Assessment” packet (see Supporting Information) in making judgments of student understanding. Prospective teachers used these learning objectives in two main ways as part of their analyses: (a) As guides in identifying relevant areas of evaluation, and (b) As an informal rubric in evaluating whether students’ responses or observed performances were correct or adequate. Consider, for example, the following excerpt:

Diana’s lab journal is evidence that she was able to identify a testable question ‘How does temperature affect crickets?’ and formulate a hypothesis ‘If crickets respond to temperature, then they will die if it is too hot or too cold.’ This is evidence she has met learning objective 4. (DM15)

In this example, a given learning objective (#4: Student will identify a testable question and formulate a hypothesis about its answer) guided the prospective teacher’s evaluation of student work by directing her attention to task-specific features related to the investigation. In these types of situations, specific performances seemed to be noticed and valued because they matched one or more learning objectives. In a fewer other cases, learning objectives were also used to make evaluative judgments of the quality of the performance. For example, consider this quote:

The idea of how they may adapt to temperature (go into a burrow if hot, seek sunlight if they are cold) is dropped. It doesn’t eliminate the value of the study, it just has moved away somewhat from the topic that was more closely aligned with the learning objectives. (DM42)

In this case, the students in the video case had first posed a question related to studying how crickets may adapt to changes in temperature, but finally decided to explore how temperature affects insect survival. Given that one of the learning objectives for the inquiry-based unit emphasized the exploration of animal behavior, what is being noticed and evaluated by the prospective teacher is a deviation from a learning expectation.

Independent Work. Many of our prospective teachers (19/43, 44.2%) noticed both student collaborations and teacher interventions and used the presence of these interactions to evaluate the extent to which assessment outcomes could be actually judged as products of self-generated, independent work. The following excerpts illustrate how noticing of independent work influenced prospective science teachers’ judgments of student understanding:

Students in this transcript reveal a moderate level of understanding because they were able to think through their hypothesis with the teacher’s help. (DM1)

At the end of the conversation Diana makes it appear that she now gets it saying, ‘All right! We rule!’ but from the evidence given, I do not think she truly understands. I think she relied on her partner to answer all of the teacher’s questions. (DM12)

In these two cases, the prospective teachers expressed reservations or uncertainty about the level of understanding of individual students because of the perceived influence of the teacher or other students. Prospective teachers who noticed the level of independence of student work often looked for evidence that indicated whether students had generated ideas or performances on their own or without excessive help from others. As was the case with Learning Objectives, noticing this type of task-general factor was often associated with judgments of the extent or quality of student understanding.
Presentation. Over a third of the student teachers (15/43, 34.9%) noticed elements related to the general format, organization, completeness, and clarity of student work, and used them to guide their evaluations. Some prospective teachers noticed the simple presence, or lack thereof, of expected elements as part of their evaluation criteria. For example,

In the poster all parts of the scientific process were accounted for: purpose, hypothesis, procedures, data tables, and analysis of results. (DM16)

The focus in this type of evaluative comment was frequently on the completeness of student work, not on the quality of the included elements. Other study participants made judgments that highlighted the value that they put on the organization of students’ presentation:

I liked how Diana clearly labeled the steps that had been taken in her experimental design. She didn’t randomly throw procedures around while coming up with them. At the end of her lab journal she included her data table that is clearly labeled with temperatures, results and equipment used. (DM6)

In cases such as this one, what was noticed was students’ abilities to present their work clearly and in an organized manner. In other instances, issues related to task completeness and organization were also used to make evaluative judgments of student understanding. For example:

The journal entry that Diana kept is evidence that she understands how to conduct scientific inquiry. Looking at Diana’s lab journal she has all of the necessary structures used in a scientific investigation. (DM38)

As illustrated by this excerpt, some of our prospective teachers made actual judgments of level of understanding based solely on the extent to which their work included all of the expected elements of an investigation or was presented in a clear and organized manner.

Task-Specific Elements

All of our prospective teachers paid attention to students’ ability to engage in one or more specific aspects of scientific inquiry as part of their evaluations (see Table 1). The different task-specific elements that were noticed can be divided into five major groups of science practices (NRC, 2007): (1) Asking questions and formulating hypothesis; (2) Designing and setting up experiments; (3) Observing and Recording, (4) Interpreting and Evaluating Evidence, and (5) Aligning Components of the Investigation. As shown below, attention to process skills related to the definition, setting up, and implementation of the investigation was more prevalent than attention to students’ abilities to building explanations and generating valid conclusions.

Asking Questions and Formulating Hypothesis. A large proportion of our study participants (86%, Table 1) paid attention to students’ abilities to formulate a research question, understand and identify experimental variables, or build testable hypotheses. In fact, these types of science process skills were among the most commonly highlighted by the student teachers, regardless of the source of evidence they selected for their analysis (see Table 2). The skills of asking questions and formulating hypotheses were more frequently cited by prospective teachers who relied on the video conversations to make judgments of student understanding (66% of all instances in which the video was used to evaluate a process skill). This could have been expected given that a majority of the conversations between students and the teacher focused on those elements. However, these skills were also predominant in
student teachers’ evaluations based on the student journal (36% of all instances) or the group poster (25% of all instances), artifacts that included evidence of students’ performance in other areas. The following excerpt is representative of the most common type of evaluative judgment made in this area by the study participants:

The transcript from Day 2 shows that the students in this group had trouble identifying what the variable was in their experiment. In the beginning of the conversation, the students don’t even seem to know what a variable of the experiment is. After help from the teacher the students are able to identify the variable, but I question if the girl really understands and would be able to identify what the variable is in another project. (DM10)

As this excerpt illustrates, many prospective teachers focused their evaluations on students’ general abilities to formulate a research question or a hypothesis, without much

Table 1

<table>
<thead>
<tr>
<th>Science Practice (Number, Percentage)</th>
<th>Inquiry Skill</th>
<th>Number, Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asking Questions and Formulating Hypothesis (37/43, 86.0%)</td>
<td>Identifying proper questions</td>
<td>8/43, 18.6%</td>
</tr>
<tr>
<td></td>
<td>Knowing what a variable is</td>
<td>27/43, 62.8%</td>
</tr>
<tr>
<td></td>
<td>Building testable hypothesis</td>
<td>29/43, 67.4%</td>
</tr>
<tr>
<td></td>
<td>Considering biological adequacy</td>
<td>10/43, 23.3%</td>
</tr>
<tr>
<td></td>
<td>of questions, variables, and hypotheses</td>
<td></td>
</tr>
<tr>
<td>Designing and Setting up Experiments (26/43, 60.5%)</td>
<td>Controlling variables</td>
<td>13/43, 30.2%</td>
</tr>
<tr>
<td></td>
<td>Having a control</td>
<td>6/43, 14.0%</td>
</tr>
<tr>
<td></td>
<td>Setting different trials</td>
<td>5/43, 11.6%</td>
</tr>
<tr>
<td></td>
<td>Setting up the experiment</td>
<td>19/43, 44.2%</td>
</tr>
<tr>
<td>Observing and Recording (16/43, 37.2%)</td>
<td>Collecting data</td>
<td>12/43, 27.9%</td>
</tr>
<tr>
<td></td>
<td>Taking notes and organizing data</td>
<td>7/43, 16.3%</td>
</tr>
<tr>
<td>Interpreting and Evaluating Evidence (14/43, 32.6%)</td>
<td>Building explanations</td>
<td>5/43, 11.6%</td>
</tr>
<tr>
<td></td>
<td>Generating adequate conclusions</td>
<td>8/43, 18.6%</td>
</tr>
<tr>
<td></td>
<td>Linking results to characteristics of life</td>
<td>3/43, 7.0%</td>
</tr>
<tr>
<td>Aligning Components (23/43, 53.5%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A teacher’s attention was counted only once regardless of the number of forms of evidence in which the teacher paid attention to a practice or skill.

Table 2

| Number of prospective teachers who used each of the listed types of assessments as sources of evidence about students’ understanding of the various science practices (N = 43) |
|-------------------------------------------------|----------|----------|
| Video Conversations | Student Journal | Group Poster |
| Asking Questions and Formulating Hypothesis | 30 | 20 | 12 |
| Designing and Setting up Experiments | 10 | 13 | 11 |
| Observing and Recording | 2 | 9 | 5 |
| Interpreting and Evaluating Evidence | 0 | 5 | 11 |
| Aligning Components | 6 | 13 | 10 |
analysis of the quality of such elements in terms of their epistemological validity (e.g., the extent to which the proposed hypothesis was testable) or biological pertinence (e.g., whether the selected independent variable could actually be expected to affect cricket behavior). Only a small proportion of the study participants (23.3%, Table 1) considered the biological adequacy of students’ ideas when attending to asking questions or formulating hypotheses, as illustrated below:

The second hypothesis shown still demonstrated that the students have not thought about the fact that crickets can change their behavior in different temperatures and do not have to die to prove it. (DM1)

In this case, the prospective teacher questioned the quality of students’ hypothesis because it failed to differentiate between effects on animal behavior, which was the target of the investigation, versus effects on animal survival of changes in the selected variable.

Designing and Setting Up Experiments. This set of science process skills was the second most frequently attended to by our prospective science teachers as part of their evaluations (60.5%, Table 1). As illustrated by the following excerpt, evaluative judgments in this area were mostly focused on students’ ability to control variables in their experiments or on students’ general ability to properly set up an experiment:

The students show this understanding by designing the experiment around one variable, making sure that they keep the rest of the experiment the same (taking temperatures for each location, keeping the number of crickets constant, etc.), and making sure they can actually perform the experiment with the available materials. (DM21)

In this evaluative comment, the prospective teacher considered issues related to experiment planning and control of variables to make judgments about student understanding of inquiry. Attention to elements related to experimental design and set-up was very similar across different types of assessment evidence (see Table 2), accounting for about 25% of all references to students’ process skills when analyzing each of the three major sources of evidence.

Observing and Recording. Over a third of the prospective science teachers (37.2%, Table 1) noticed students’ abilities to collect and organize experimental data. A large fraction (56.2%, Table 2) of evaluative judgments in this area was made based on the analysis of evidence presented in the form of a student journal entry. Close to two thirds of these types of evaluative comments referred to students’ ability to collect data in an organized manner:

She seems to have a good grasp on how to collect the data in an organized table. (DM5)

A smaller proportion of study participants focused on students’ ability to collect adequate data to prove or disprove the hypothesis. The following excerpt illustrates this other type of analysis:

All of the data on the table is relevant to the hypothesis and test. They did not include any impertinent observations or data on the table. This shows that they understand the difference between serendipitous observations and deliberate observations. (DM16)

Here, the evaluation of data adequacy is based on a judgment of alignment between three different elements of the investigation: hypothesis, experimental design, and data collected.

Journal of Research in Science Teaching
Interpreting and Evaluating Evidence. Some prospective science teachers (32.6%, Table 1) noticed aspects of students’ ability to build explanations and generate adequate conclusions as part of their evaluations. However, this group of practices was the least prevalent in all of the coded judgments of student understanding and performance. As illustrated in the following excerpt, the few participants who attended to elements in this category noticed the lack of data analysis, explanations, or conclusions in the student journal as part of their evaluations:

They show a results table, but no verbal explanation of the results and no written conclusion either. There needs to be an explanation of their results, and an interpretation of the results to indicate a specific conclusion based on their findings. (DM19)

A smaller proportion of the prospective teachers (7%, Table 1) analyzed the biological ideas presented or discussed by students as part of the group poster. The following excerpt illustrates this type of analysis:

The crickets were more attracted to the green construction paper on the side of the box, and did not stay on the brown ground. It is an extremely interesting observation that didn’t make a connection to any of the characteristics of living organisms. I think it was a missed opportunity for the students to use knowledge they did have to explain the phenomenon that they were seeing. (DM32)

In this case, the focus of the evaluation is on students’ failing to build connections between their observations and assumed prior knowledge about the characteristics of living organisms. In general, these types of evaluative judgments did not describe in detail the specific biological ideas that the students failed to connect or understand.

Aligning Components of the Investigation. Over half of the prospective science teachers (53.5%, Table 1) attended to issues of coherence or alignment between different components of students’ investigations. For example:

There was a large disconnect between the hypothesis/purpose of their experiment and the procedure and data. Their question is in regards to individual activity that is independent from one another however they measured the activity of all the crickets for different activity-tasks. (DM28)

This excerpt illustrates evaluation of alignment between three components of the investigation: hypothesis, experimental procedure, and data collected. Analysis of alignment issues was most common in evaluations based on the student journal and the group poster (see Table 2). Our analysis revealed that 35.5% of the evaluative judgments made about alignment referred to issues of alignment between components within the category Asking Questions and Formulating Hypothesis (e.g., whether the hypothesis was consistent with the independent variable selected or whether the research question was aligned with the hypothesis). Another large proportion of these types of judgments (45.2%) were based on the analysis of the consistency between components in the latter category and components within a different science practice, mainly the experimental procedure that the students’ followed (16.1%) or the type of data that they collected (19.1%). A small proportion of these evaluations (12.9%) considered issues of alignment across more than two major groups of science practices.

Overall, 18.6% (8/43) prospective teachers in our sample noticed elements belonging to only one of the five major groups of science practices used to categorize results in our study. 62.5% of these participants paid attention to elements in the Asking Questions and

Journal of Research in Science Teaching
Formulating Hypothesis category and 37.5% noticed elements in the Designing and Setting Up Experiments group. A similar number of prospective teachers (9/43, 20.9%) noticed elements from two different sets of science practices while twice as many study subjects (18/43, 41.9%) attended to elements belonging to three different groups. 88.9% of the prospective teachers who noticed elements in two or three categories of science practices attended to elements in the Asking Questions and Formulating Hypothesis category while 51.9% noticed elements in both the Asking Questions and Formulating Hypothesis and the Designing and Setting Up Experiments groups. There were a few prospective students who considered elements in four (3/43, 7%) or five (5/43, 11.6%) of the major groups of science practices in their evaluation of student work. These results indicate that although most prospective teachers noticed elements in more than two categories, many of them exhibited preferential attention for science process skills associated with the setting up and design of an investigation.

Discussion

Three major assertions can be developed from the analysis of our findings:

Assertion #1: Prospective teachers paid attention to both task-general and task-specific elements in the evaluation of student work in the context of an inquiry-based learning experience. Task-general elements often served to guide or qualify the specific inquiry skills that were evaluated.

Assertion #2: Prospective teachers paid attention to several task-specific elements associated with students’ performance in different parts of the investigation. However, their analysis was often limited to highlighting the presence of expected components without further evaluation of the actual ideas expressed by the students.

Assertion #3: Prospective teachers preferentially attended to aspects related to the design of an investigation. They paid insufficient attention to science practices associated with interpreting and evaluating the results of an investigation.

These assertions reveal strengths and weaknesses in prospective teachers’ assessment reasoning that can serve as guides in the development of interventions to challenge and enrich their conceptions of student understanding in the context of inquiry-based learning.

Prospective teachers involved in our study were undergraduate students who had completed at least five courses on teaching and learning science in secondary school classrooms, as well as close to 200 hours of internships in such settings. Our findings suggest that many of them had begun to notice important elements in the evaluation of student understanding in the context of inquiry. Some of these elements were task-general in character, such as looking for alignment between students’ performance and stated learning objectives, searching for evidence of students’ ability to complete a task in an independent manner, or paying attention to the organization of students’ ideas. The identification of these types of task-general factors is important because they direct prospective teachers’ attention to specific elements of students’ performance and influence judgments of student understanding. Our results suggest that general knowledge and beliefs about teaching and learning, such as designing lessons according to learning objectives, influence what teachers notice and how they frame the evaluation of student work.

Study participants also noticed important task-specific elements in the evaluation of student performance during an inquiry-based learning experience. In particular, many of prospective teachers paid attention to more than one important component of the investigation and to the alignment between them. However, our results also indicate that prospective
teachers’ assessment reasoning often displayed preferential attention for the form (i.e., presence of inquiry components) over the substance (i.e., specific biological ideas) of student work, and was more focused on the evaluation of students ability to design and set up experiments than on analyzing results and generating plausible conclusions based on the data and their scientific knowledge. This type of preferential attention was independent of the nature and content of the artifacts used to evaluate student understanding (i.e., whether the artifact included evidence of some or all of the components of the investigation). This pattern of noticing was also characteristic of the judgments made by study participants about aligned or misaligned components of the investigation.

The above results are, to some extent, similar to those reported by Ruiz-Primo and Furtak (2007) in their analysis of middle school science teachers’ informal formative assessment practices. In that case, researchers noticed a tendency for teachers to ask more eliciting questions related to procedures and observations than to the formulation of explanations and arguments. However, we acknowledge that generalizations about prospective teachers’ preferences in noticing during classroom inquiry lessons or units cannot be drawn from our exploratory study. The specific nature of the selected assessment task may have constrained the types of elements noticed by our prospective teachers during their evaluations. The use of one assessment task completed by the teachers and one classroom inquiry unit to focus their attention on in completing the assessment limits generalizability, but nonetheless raises interesting questions for future exploration of teachers’ noticing.

The reasons for prospective teachers’ preferential attention from some elements of student work over others may be diverse, and more research will be needed to elucidate them. However, one may suggest plausible explanations based on related work on teacher thinking. On the one hand, it is possible that some participants’ assessment reasoning might had been constrained by an incomplete or a naïve view of inquiry. Recent research on secondary school science teachers’ (Demir & Abell, 2010) and college science instructors’ (Brown, Abell, Demir, & Schmidt, 2006) views of inquiry indicates that these practitioners frequently fail to include essential components of an investigation, such as evidence, explanation, justification and communications, in their meaning of inquiry. When talking about inquiry, the participants in the aforementioned research studies emphasized that student-generated questions and carrying independent research were the foremost features of inquiry. Ascription to this view of inquiry may explain both our participants’ preferential attention to the initial phases of inquiry, as well as their focus on independent student work in judging the extent or depth of student understanding.

An incomplete or naïve view of inquiry may also be responsible for the lack of attention to and evaluation of the biological adequacy of students’ ideas. If our prospective science teachers believed that the main goal of inquiry in the science classroom is to help students learn “how to think” or “how to solve problems” (Brown et al., 2006), content understanding might had considered as a secondary goal during inquiry-based activities. Windschitl (2004) has suggested that prospective science teachers tend to subscribe to a “folk” theory of inquiry in which this process is mainly seen as a technical procedure rather than as a theoretically grounded exploration. From this perspective, learning to inquire in science may be conceived as the process of learning a set of discrete process skills rather than as developing the ability to test, develop, or compare models framed by scientific theory. This artificial separation between science content and practices has also been noted by Roth and Garnier (2007) in their analysis of TIMSS video data from science classrooms.

However, it is also possible that the patterns of noticing elicited by our analysis might have been the result of how some of our study participants framed the task of evaluating
student understanding in the context of an extended investigation, rather than by their incomplete or preconceived views of inquiry. Recent studies on the nature of novice (Levin et al., 2009) and more experienced (Tang et al., 2010) science teachers’ attention in the classroom suggests that science teachers have abilities for attending to student thinking, but what they notice depends on how they frame what they are doing. Most teachers work in professional environments in which powerful contextual factors focus their attention on classroom routines, students’ behaviors, and curriculum fidelity, rather than on student thinking. In the case of the prospective teachers involved in our study, it is difficult to evaluate the extent to which school culture might have shaped their approach to evaluating student work in the context of inquiry. Although our study participants were several weeks into their student teaching experience, the task used in our study was part of an evaluation associated with the university component of their teacher preparation program. Under those circumstances one could have expected our prospective teachers to frame the task as a requirement of one of their science education courses in which analysis of student understanding of science content was highly valued. The observed lack of attention to the substance of students’ ideas may thus be indicative of the failure of such courses in promoting the type of thinking that is desired.

We also acknowledge that the specific structure of the task used to build our study, as well as the nature of the evidence made available to the prospective teachers to complete the assignment, may have influenced the type of noticing that we observed. For example, as part of our “Decision-Making Assessment” task, prospective teachers were asked to evaluate the teacher’s planning decisions based on the analysis of the learning objectives included in the provided artifacts. This activity could have primed prospective teachers to use learning objectives as lenses to guide their noticing. Additionally, the video tape of the inquiry-based unit given to our study participants only presented small segments of the entire activity, several of which involved students working in the initial phases of their investigation (i.e., defining a testable question, identifying a variable, posing an initial hypothesis). It could be possible that the nature of the video case constrained what was noticed by over-representing some aspects of students’ work. Ultimately, prospective teachers had access to a limited number of artifacts presented in a specific format (e.g., excerpt of lab notebook of an individual student, copy of a final group poster) that may have facilitated their noticing of some aspects of student understanding while constraining others. Nevertheless, the consistency in the responses written by prospective teachers who based their evaluations on different types of artifacts suggests that the observed patterns of noticing were not fully determined by the type of evidence provided to them.

Implications

The findings from our study pose a challenge for preparing prospective teachers to successfully teach in alignment with the new Framework for K-12 Science Education (NRC, 2011) and the Next Generation Science Standards. These documents call for redirecting teaching about scientific inquiry, away from experiencing inquiry as a set of process steps and toward experiencing inquiry as the development of arguments, explanations, and models built from critical interpretation and evaluation of evidence. Yet, these are exactly the often unrealized elements of inquiry to which the prospective teachers in this study gave least attention. Rather, the prospective teachers noticed students’ abilities or inabilities to complete the initial phases of an investigation, which may be indicative of a “process” orientation toward teaching in which the development of science process skills such as posing questions and designing experiments is viewed as the end or goal of instruction, instead of the means to an end.
for developing conceptual understanding (Millar & Driver, 1987; Talanquer, Novodvorsky, & Tomanek, 2010).

So, how might this challenge be addressed by science teacher preparation programs? Perhaps a first step might be providing opportunities for prospective teachers to identify and reflect upon both the task-general and the task-specific elements that guide their noticing and evaluation when asked to interpret students’ understanding of inquiry. In this regard, engaging prospective teachers in the adaptation and development of reform-based inquiry curriculum may be effective in promoting teacher learning and reflection about inquiry instruction (Davis & Krajcik, 2005). Other researchers have highlighted the importance of deconstructing and collectively analyzing models of inquiry implemented in their science teacher preparation courses (Schwarz, 2009). Creating opportunities for prospective teachers to engage in either inquiry experiences guided by theoretical models or well-supervised authentic science practices (Windschitl, 2003, 2004) could also be beneficial in helping prospective teachers develop inquiry teaching views compatible with the Next Generation Science Standards.

Efforts to develop prospective science teachers’ views of inquiry may also be useful in revising the teachers’ patterns of noticing and the reasoning used to assess students’ understanding of inquiry. In particular, engaging prospective teachers in extended apprenticeships in the epistemic discourse of science, based on the Model-Based Inquiry approach to investigative science (Windschitl, Thompson, & Braaten, 2008a, 2008b), seems promising. Although not without challenges, most prospective science teachers involved in the Model-Based Inquiry learning environments came to reconceptualize the roles of models, theory, evidence, and argument in scientific inquiry. The reframing of science practices supported a shift in prospective teachers’ goals for scientific investigation, from testing predictions and proving hypotheses, to testing and revising theoretically grounded explanatory models. We suspect that a learning environment similar to the Model-Based Inquiry approach may also support a shift in what prospective teachers notice when assessing students’ understanding of inquiry.

Finally, the challenge might be addressed by designing and implementing activities in our science teaching classes that provide prospective teachers with increased practice in attending to student thinking in the context of inquiry. Such learning opportunities may help prospective teachers frame teaching and assessment in terms of making sense of students’ ideas (Levin et al., 2009). Our results suggest that prospective teachers overlook relevant evidence or fail to analyze it in depth when making judgments about students’ understanding. The lack of attention to what students actually say, write, or do, or the inability to make productive sense of it, seriously limits teachers’ abilities to support and foster student learning (Otero & Nathan, 2008). Science teachers not only need to learn to respond to students’ ideas. They must also learn what is most critical to notice and respond to. From our perspective, assessing student understanding by attending to student thinking is a critical centerpiece of teacher learning in a science teacher preparation program.

References


Journal of Research in Science Teaching


Journal of Research in Science Teaching


Journal of Research in Science Teaching


