

Assessing Science as Inquiry in the Classroom

**Pamela Van Scotter, BSCS (Biological Sciences Curriculum Study)
K. David Pinkerton, BSCS**

IO

Inquiry is essential to scientific endeavors. Consequently, it is essential that students understand and use it. Several previous chapters of this book have described how questions, explanations, argumentation, and interpretation are aspects of inquiry that should be part of school science instruction. Other chapters have provided glimpses of what inquiry might look like in specific contexts, such as in a chemistry class or an Earth science class. Now that some clear models for inquiry-based instruction have been established, this chapter will describe some ideas about assessing for understandings and abilities related to inquiry.

Teachers need to have strong ideas about what will serve as evidence of students' understanding of inquiry before they begin teaching (Wiggins and McTighe 2005). Once teachers know what will serve as evidence for understanding, then sequencing effective learning experiences can be more targeted and more clearly focused.

Assessing Inquiry and Links to How Students Learn

When assessing students' understandings of and abilities in regard to inquiry, teachers should ensure that the assessments align with instruction and with the current findings about how students learn. Using Table 10.1, consider some characteristics of inquiry assessments that are consistent with the research of Bransford, Brown, and Cocking (2000).

TABLE 10.1. ASSESSING FOR INQUIRY CONSISTENT WITH KEY FINDINGS ABOUT LEARNING

Key Findings From <i>How People Learn</i>*	Implications for Assessing Science as Inquiry
1. Learners come to the classroom with conceptions about the world around them.	<ul style="list-style-type: none"> • Assessments should provide opportunities for students to demonstrate what they know and what they don't know about scientifically testable questions, designing investigations, using evidence to develop explanations, and communicating findings to others.
2. Learners need a deep foundation of conceptual knowledge upon which to build improved understandings.	<ul style="list-style-type: none"> • Assessments need to be conceptually coherent and connected to the students' current foundation of understandings of and abilities in regard to science as inquiry. • Assessments need to be ongoing so as to continually add layers to the foundation of knowledge and abilities related to inquiry. • Assessments need to be balanced to provide a complete picture of what students know.
3. Learners need opportunities to monitor their own learning.	<ul style="list-style-type: none"> • Assessments need to provide students with opportunities to assess themselves in ways that help them monitor how strong their understanding is and at what point it begins to break down. • Assessments need to be relevant and compelling so that students will see the value in monitoring their understanding.

*The findings in column one are adapted from Bransford, J. D., A. L. Brown, and R. R. Cocking, eds. 2000. *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press, pp. 14–18.

Balanced Assessments of Inquiry

Teachers should ensure that assessments of inquiry are balanced and authentic. By *balanced*, we mean that teachers should assess the many ways and the many points at which students represent what they know or are in the process of learning about science as inquiry. By *authentic*, we mean that the assessments should reflect the nature of science and the practice of the scientific endeavor.

Balanced Assessments Include a Range of Assessment Types

As a first step, consider the types of assessments that, taken together, make up a balanced approach. Using the principles of backward design (Wiggins and McTighe 2005), these assessment types follow from the desired outcomes of inquiry (NRC 1996), as shown in Table 10.2.

TABLE 10.2. OUTCOMES OF INQUIRY ALIGNED WITH A RANGE OF ASSESSMENT TYPES

Targeted Outcome of Inquiry	Assessment Type	How Outcome Aligns With Assessment Type
Understanding about inquiry	Endpoint (summative, point-in-time, static, delayed feedback)	Inquiry is a body of knowledge that can be tested objectively after instruction.
Ability to perform inquiry	Dynamic (formative, ongoing feedback, “in the moment”)	Inquiry is a process that can be monitored and evaluated as students perform it.
Conceptual understanding	Conceptual framework (cognitive associations, mental structures, conceptual connections)	Inquiry produces changes in students’ conceptual frameworks, moving them toward more expert-like thinking.

Endpoint assessments evaluate what students know at a single point in time, usually after instruction. For example, tests and quizzes give teachers information on how students respond to classroom activities. Typically, endpoint assessments require students to interact with nonchanging (static) prompts, such as multiple-choice or free-response questions. Feedback from teacher to student is delayed. Endpoint assessments make it difficult for teachers to know exactly

what students are thinking, especially when written responses are sparse or difficult to interpret. These assessments do provide important objective documentation of a student's abilities at a moment in time, however.

Dynamic assessments determine how students respond to changing (dynamic) prompts. For example, teachers use questions in conversations with students to pinpoint what they can answer on their own and what they need help to answer. The fluid nature of teacher-student, real-time feedback provides information for ongoing learning (see Vygotsky 1962/1986). The same type of back-and-forth interactions occur among students during investigations (Hake 1992). This is especially true when teachers orchestrate conversational prompts to probe students' prior conceptions, to link evidence and interpretation, or to construct meaning through argumentation and inference. Student notebooks also provide a record of these "in the moment" interactions. Teachers can use notebooks to assess the ongoing aspects of student learning (as described in Biological Sciences Curriculum Study [BSCS] 1994).

Conceptual framework assessments give teachers insight into the way students store, retrieve, and associate knowledge. For example, the concept maps and graphic organizers of experts in a field are consistently different from novices (see Novak 1990). The association and hierarchy of concepts form the basis of mental structures used to do well in school. Cognitive framework assessments inform teachers about the effect their teaching has on those structures.

Designing Balanced Assessments

One way to ensure a balanced approach to assessment is to design a scoring scheme or rubric that reflects the desired balance among endpoint, dynamic, and conceptual framework assessments. This way, teachers have data so they can effectively triangulate what students know and can do. As a result, teachers can improve their ability to individualize instruction.

For example, imagine a unit on designing, conducting, and communicating results of a scientific investigation. How do you construct a balanced rubric for this unit on inquiry? Naturally, assessing inquiry works best and most efficiently within a context, namely, specific science content. For this example on assessing inquiry, consider the following focus question posed to students: "What affects the period of a pendulum and how?" Note that in a "full" inquiry students would formulate their own questions, but in many actual school circumstances, time constraints do not allow full inquiry in every activity. Nonetheless, each aspect of inquiry can be assessed, documented, and used to guide instruction.

Before students begin their investigation, the teacher communicates all assessment criteria in the form of a scoring rubric. The rubric tells students ahead of time what is important and what a high-quality performance looks like. Rubrics accomplish this, in part, by reflecting the intended balance of assessment types. The inherent balance structured into the rubric sends a clear message to students about what is essential to learn. In other words, if you say balance is important, then it should be assessed in student performance.

How do you build a balanced rubric? A rubric *design matrix* can help ensure a balanced assessment. Think of a design matrix as systematic brainstorming. That is, creatively generate ideas for assessments by thinking of examples from specific categories of assessment. In the design matrix, list the concrete tasks to be assessed in the task column. Generally, only one task for each category of assessment is necessary.

Next, form questions that a specific assessment type would tend to answer. Note that students will not have to answer all questions from each category of assessment. The intent at this stage is to generate as many questions per category as possible, as in a brainstorming activity. Eventually, convert some of these questions into concrete, performance-level indicators associated with specific tasks. These performance criteria show up in the final rubric.

Since the assessment categories represent, in sum, a balanced assessment, then chances are good that the final rubric will also be balanced. When the rubric is balanced and students use the rubric to guide their work, chances are that the work students do will also be balanced. Table 10.3 is a completed rubric design matrix for the pendulum-period activity.

After filling each cell in the design matrix, use the completed matrix to decide the best match between task and assessment type. For example, in Table 10.2, the individual multiple-choice test best aligns with endpoint assessments. Eventually, each task is assessed by one type of assessment. Of course, there is overlap, but you decide what the final division will be. A balanced assessment results from spreading the types of assessment among the tasks on the rubric and giving significant weight to each assessment type in scoring. Finally, write performance-level indicators for each task and decide how to weight each task for scoring purposes.

The rubric in Table 10.4 results from the design matrix just discussed. In this rubric, teachers use the concept map task to assess conceptual framework knowledge, the multiple-choice test as an endpoint assessment, and documentation in students' notebooks (along with teacher observations during

TABLE 10.3. EXAMPLE OF A DESIGN MATRIX THAT HELPS ENSURE A BALANCED ASSESSMENT FOR A GIVEN TASK

Task	Endpoint	Dynamic	Conceptual Framework
Generate, complete, and evaluate the results of a systematic procedure to answer the focus question.	<ul style="list-style-type: none"> • Does the student produce a procedure? • Does the student carry out the procedure? • Does the student evaluate the results? 	<ul style="list-style-type: none"> • Are there various drafts of the procedure in the student's notebook? • Are there notes from discussions during design iterations? • Do you see evidence of ongoing adjustments to the design, based on feedback from others? From the phenomenon? • Is there evidence of an emerging conclusion based on ongoing evaluation? 	<ul style="list-style-type: none"> • Does the student use the language of inquiry properly? • Does the student use the focus question to make decisions about steps in the design? • Does the student construct conclusions based on evidence? Are they logical and consistent?
Take individual multiple-choice test regarding essential understandings about inquiry.	<ul style="list-style-type: none"> • What is the raw percent correct? 	<ul style="list-style-type: none"> • How much time was required? • Did the student request clarification or elaboration during test? • What appears on the scratch paper? 	<ul style="list-style-type: none"> • What type of questions were missed most often (conceptual, algorithmic, or recall)? • How does distribution of missed items compare to highest-performing students?

Table 10.3 continues on next page.

Table 10.3 continued.

Task	Endpoint	Dynamic	Conceptual Framework
Review and analyze a research article about the essential understandings of scientific investigations.	<ul style="list-style-type: none"> • Are there notes in the student’s notebook? • Did the student list key words in the notebook? • Can the student write an effective summary paragraph analyzing the research and the findings? 	<ul style="list-style-type: none"> • Do students use literacy strategies during reading? • Do students debrief with peers about the article? • Are there notes from peer debriefings that indicate thoughtful discussion? 	<ul style="list-style-type: none"> • Can the student generate a concept map using essential terms of the article? • Does concept map demonstrate coherent principles and conceptual connections? • Is the extent and sense of relatedness among concepts similar to what teacher-experts might produce?

lab) to focus on dynamic assessment. The end product is a balanced assessment from which teachers produce a balanced learning environment—one most capable of enhancing achievement for a broad diversity of learners.

Authentic Assessment: Capturing Information About Students’ Practice of Science

To more fully assess and monitor students’ abilities of inquiry in the classroom, teachers and students can use the same type of rubric presented in Table 10.4 but in this case—Table 10.5 on page 116—designed specifically to focus on these abilities of inquiry (NRC 2000). So, for example, throughout a year or a semester, students can set goals for themselves and monitor their progress toward these goals. Teachers also can use such a rubric to formally assess students’ performances on a specific investigation, as well as to

TABLE 10.4. EXAMPLE OF A BALANCED SCORING RUBRIC



Weight	Task	3	2	1
40%	Design, conduct, and evaluate the results of an investigation to answer the focus question.	Exhibits detailed ongoing progress of design process by documenting the development of procedures, the responses to peer feedback, and the generation of conclusions in notebook.	Documents and exhibits an understanding of the final design only, not ideas and drafts of ongoing design development.	Records data from investigation with little reference to ongoing procedural adjustments and developing interpretations of data.
30%	Take individual multiple-choice test regarding essential understandings about inquiry.	Selects between 81% and 100% of the multiple-choice answers correctly.	Selects between 61% and 80% of the multiple-choice answers correctly.	Selects between 0% and 60% of the multiple-choice answers correctly.
30%	Create a concept map, demonstrating a rigorous review and analysis of an article about the essential understandings of inquiry.	Designs, generates, and explains a concept map containing key ideas related to understandings of inquiry, which reflect article's conceptual hierarchy, connectedness, and application.	Constructs concept map with key concept terms, but does not show and explain the hierarchy, connectedness, or application among those terms from the article.	Produces a physical arrangement of terms from the article that has little discernible rationale and does not reflect hierarchy, connectedness, or application of the key understandings of inquiry.

Note: The balanced scoring rubric weights each task sufficiently to communicate the importance of all outcomes to students.

discern patterns of strengths or weaknesses across the class. These assessment practices align with the goals for assessment outlined in Table 10.1. This type of information helps a teacher structure the subsequent learning experiences to target the type of practice and feedback students need most.

The Influence of Instructional Models

Constructivist instructional models represent an example of a family of instructional models that, among other things, foster balanced and authentic assessment. Such models orchestrate learning for students in a way that supports enduring understandings and provides a framework for constructing knowledge about scientific inquiry (Bybee 1997). Specific examples can help demonstrate the ways in which stages of an instructional model are linked to different types of assessment and how these assessments are also learning opportunities for students.

In the invitation or engagement phase of an instructional model, as well as during the discovery or exploration phase, opportunities for dynamic assessments of abilities and understandings of inquiry are introduced. For example, imagine that Ms. Washington, a middle school science teacher, had assigned her students the task of designing an investigation as part of ongoing class activities. She can now become part of the assessment process through dialogue with students, especially in the form of questions. She can monitor the design process by observing student interactions, reviewing student notebooks, and evaluating the activity results.

As part of monitoring the design process, she would involve herself in informal feedback sessions with students. Then she would use the trajectory of these interactions to assess the dynamic (ongoing) growth in a student's inquiry abilities. Examine the following dynamic assessment version of Ms. Washington's design task and look for evidence of in-the-moment learning.

S: Ms. Washington, what do we do next? You didn't give us a worksheet with the steps.

T: Does the focus question give you any ideas?

S: It just asks what happens to volume if temperature changes. It doesn't say what to do.

T: How did you know to *change* temperature? The question didn't have that in there.

TABLE 10.5. SAMPLE COMPLETED RUBRIC



Higher Levels of Performance

Weight	Features of Inquiry*	3	2	1
15%	Engages in scientifically oriented questions.	Consistently recognizes and poses appropriate questions that can be answered by science.	Is sometimes able to recognize and pose appropriate questions.	Rarely distinguishes between questions that can and cannot be answered by science.
25%	Gives priority to evidence as he or she explores answers to questions.	Consistently understands and determines what appropriate evidence is and designs investigations to collect it.	Is sometimes able to determine what evidence is needed to answer simple questions and design simple investigations.	Rarely is able to determine what evidence is needed to answer a question or to design an investigation to collect it.
25%	Formulates explanations based on evidence.	Evaluates the evidence critically and consistently uses appropriate evidence to formulate explanations.	Formulates intermediate explanations using mostly appropriate evidence.	Formulates weak explanations, often using unrelated evidence.
25%	Connects his or her new explanations to a rich base of current knowledge.	Seeks out additional information and connects it appropriately to current explanations.	Seeks out other information but is not able to connect it to current explanation.	Rarely connects information from other sources.
10%	Communicates and justifies the explanations he or she constructs.	Develops and communicates a logically consistent explanation.	Develops an intermediate explanation that includes some inconsistencies or uses only partial evidence.	Develops a weak explanation that is logically inconsistent and uses little evidence.

Note: This rubric represents a tool for both students and teachers as they observe the students' inquiry abilities improve during the year.

*The Features of Inquiry in column two are adapted from National Research Council (NRC). 2000. *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press.

- S: How else are we supposed to make the volume do something?
- T: Good thinking, Eddie, you have to change the temperature. Write down how you're going to do that in steps that other students could follow.
- S: Is that all we have to do?
- T: Well, what else have you done when you have designed investigations in class?
- S: Let's see. We've been careful to change the thing we want to study and keep the other things the same.
- T: OK, now you're putting some good ideas together. Try putting those ideas in your design. Then call me over and we'll talk some more.

Ms. Washington didn't accept Eddie's initial knowledge position as an accurate reflection of his true ability. So she thought of questions to probe Eddie's deeper understanding. In Eddie's first response, she noticed how Eddie had inferred the need for a change in temperature from the focus question. Then she prompted specific actions (writing a design and procedure) that caused Eddie to formalize and record his thinking. Finally, she foreshadowed the need for continued feedback as she and Eddie worked toward a finished design.

At the end of an explanation-type activity, a teacher might ask, "Which of the following is a question that can be answered by science?" If a student chooses an incorrect response such as, "Is it right to euthanize unwanted dogs?" the response informs the teacher about the student's specific understandings regarding inquiry. This type of feedback often comes at the end of an activity when students have been developing their understandings and explanations, and so this is an example endpoint assessment.

Application and evaluation-type lessons can be used to assess changes in students' conceptual frameworks. We know that this is a critical first step in learning (see Bransford, Brown, and Cocking 2000). For example, to engage students' prior knowledge about using evidence to develop explanations, the teacher might have the students construct a concept map using their current understanding of the importance of evidence. After progressing through conceptually coherent, carefully sequenced activities, students would be asked to construct another concept map from the same concept words. By comparing the before and after

maps to an “expert” map, teachers can validly assess changes in the conceptual framework of their students (Pinkerton 1998). Other forms of assessing aspects of changes in students’ conceptual framework include performance on conceptually oriented tests, generating graphic organizers that convey a conceptual hierarchy and connectiveness, and solving open-ended problems, such as those typically found in authentic assessments.

Student Self-Assessment: When Inquiry Gets Personal

Balanced and authentic approaches to assessing inquiry teach students about multiple sources of feedback, not just feedback from the teacher. One of the most important sources of feedback is when students monitor their own thinking (metacognition) (Bransford, Brown, and Cocking 2000). This monitoring reflects the important goals for assessment outlined in Table 10.1 that link assessment to what we know about learning. Self-assessment involves an ongoing and iterative interaction with the prompt at hand, be it a reading, an investigation, or an open-ended project. Students use the information they obtain from self-monitoring to assess their understanding and adjust their learning paths. The use of a rubric such as the ones in Tables 10.4 and 10.5 are useful tools for students in this process. In many ways, the active meta “dialogue” required by student self-assessment is at the heart of effective inquiry.

Using dynamic and conceptual framework assessments in addition to endpoint assessments teaches students self-assessment skills. For example, consider what it takes to teach students how to learn from mistakes. It requires explicit instruction on how to pinpoint, articulate, and remedy mistakes on content-related activities. In effect, students who learn from mistakes learn to monitor their ongoing thinking and use it to plan actions and solve problems. Used with unit tests, learning from mistakes through self-assessment and peer dialogue shifts endpoint assessments toward dynamic assessments and fosters greater balance and higher achievement (Pinkerton 2005).

When students learn and practice self-assessment skills, they move toward greater intellectual independence. They tend not to require as much external feedback from teachers and can chart their own way through an investigation. In turn, teachers can allot more of their time to those students who still are struggling, thus spending time where it is most needed. In effect, teaching self-assessment skills shrinks the class size by increasing the number of students who can apply the outcomes of inquiry to doing the business of school.

Conclusion

The message is clear—the ability of students to understand and to do inquiry is essential. Developing inquiry-based lessons and assessing for understandings and abilities of inquiry go hand in hand. In fact, well-designed inquiry-based lessons result from well-designed assessments—ones focused on the important features of inquiry. This backward design approach begins with the end (the assessment), which in turn, provides a specific target for the learning sequence that builds toward the assessment task. That is, the process begins with what is important for students to know (the essential skills and understandings of inquiry) and ends up with an effective learning environment (the day-to-day experiences that shape what students know and understand).

Assessments of inquiry should align with what we know about learning and should be balanced and authentic. That is, they should include all the important features of inquiry, not just those easy to assess. This approach helps ensure that all students acquire the knowledge of and skills regarding scientific inquiry considered important. In turn, this foundation of knowledge should help students participate more effectively in an increasingly complex world.

