Introduction

> When teachers are asked how they assess their students, they typically talk about tests, examinations, quizzes, and other formal methods. When they are asked how they know whether their students have learned what they have taught, the answers are very different.
>
> —Dylan Wiliam, *Assessing Science Learning* (2008, p. 6)

Our districtwide K–12 science team came back from a National Science Teachers Association conference last spring all fired up. We had gone to a session on formative assessment and learned about the *Uncovering Student Ideas in Science* series of assessment probes and a variety of science formative assessment classroom techniques (FACTs) we could use with the probes. The presenter started the session by asking us to call out the first word that came to mind when we heard the word *assessment*. In unison, most of the people in the room called out "testing!" As the presenter then pointed out, *assessment,* and particularly formative assessment, is not necessarily about testing; it is about what you can do to improve learning and, ultimately, get better test results.

During the session, we, as learners, used the probes and the FACTs and came to realize how powerful they are. We had heard the phrase *assessment* ***for*** *learning* before, and for

Introduction

the first time, we felt we really knew what it meant. The speaker talked about the importance of using assessments like the probes and FACTs to create a balanced system of assessment that provides useful information at the beginning of instruction to promote thinking and inform instruction. This "front-end" assessment leads to better results at the "back end," when students are tested on what they have learned.

This approach to assessment started to make sense to us. All the concepts and ideas from the practice tests our students were taking to prepare for the state assessment were quickly forgotten by the students, even when it came time to take the test. We knew drilling with sample test questions wasn't the best solution for raising test scores, but we had always hoped it would help a little. We now realized that maybe formative assessment was what we needed to do more of at the beginning and throughout a lesson so that students—rather than memorize a lot of discrete facts—would have an opportunity to confront and work through their ideas before taking a test. After all, science is different from other subjects—it's about ideas that explain the natural world and processes that help us make sense of that world.

Many of the people in the audience had been using the probes and shared their stories. We connected with several folks from districts like ours and got all kinds of good ideas about how to use these tools to improve student learning and teacher practice. Before we left the conference we went to the NSTA bookstore and bought copies of the *Uncovering Student Ideas in Science* series. We started reading them on the plane ride home. We couldn't stop talking about them and our ideas for using them! We felt we had finally found a solution to the struggles we face in our district in balancing accountability and reporting with students' opportunity to learn.

Teaching and assessing for understanding aren't about more teaching, more materials, and more testing; they are about more opportunities to learn by promoting thinking and bringing students' ideas to the surface. Here was a set of assessments already developed for us to use with links to the key ideas in the standards, descriptions of the research that the probe was based on, and suggestions for things to do in the classroom to help students learn. These assessments would save us months of work that might have been spent developing our own formative assessments.

When we got back to our district we shared what we had learned in our grade-level teams with teachers and administrators. We argued over what *formative assessment* really meant and whether our current practices were consistent with what the research describes as good assessment practice. We all decided we wanted to move beyond practice for test taking and deadly drill sessions. Our superintendent surprised us with her enthusiasm

and her offer to provide funding to support stipends and copies of the series for after-school professional learning communities. These groups would come together and study this new assessment technique, try it out in the classroom, examine student results, and report back.

After several months of exploring the theory and practice of formative assessment in collaborative groups, the broad consensus of our learning communities was that formative assessment worked! Students were more engaged in learning science, they began to write more extensively and to converse scientifically about their ideas, and they took more ownership in the learning process. Their explanations became much richer and we knew better how to then tailor learning to address the students where they were, rather than where our textbook and pacing guide said they should be. Our classroom questioning had changed from an ongoing monologue back and forth between the teacher and students to one in which there was rich dialogue among students working together with guidance from us in resolving their ideas. All students were involved and felt safe to share their thinking.

As we tried out formative assessment, we also looked at our summative assessments. We quickly found that released items from prior state assessments, our district-developed standardized tests, and even our own classroom tests did not give us the kind of data we needed to know exactly what the students' learning problems were. We knew our students weren't doing well in some areas such as matter and energy interactions or Earth systems, but we didn't know exactly what the learning problems were until we used the formative assessment probes as well as other new teaching strategies that probe students' thinking.

We decided to match up the probes with the assessment reporting categories aligned to our state standards, administer the probes across grade levels, analyze the results, and match the results to our district test data. Lo and behold, we found that the problem areas for tenth graders weren't much different from the problems of our middle school and elementary students.

Good data about student learning are at the core of improvement in student achievement. We began to use the Collaborative Inquiry into Examining Student Thinking (CIEST) protocol to look at student work from the probes (Mundry, Keeley, and Landel 2009). The probes revealed not only information we could use to inform our instruction but also a lot about the gaps in our K–12 curriculum that were affecting learning from one grade level to the next. No wonder certain problems continued after the fourth-grade test and the eighth-grade test. As a district, we had never collected the rich kind of formative data that could be used to pinpoint what the learning problems were, how they originated,

Introduction

and why they persisted from one grade to the next. Instead of looking at test scores—a single snapshot in time—and saying we needed to reteach the same material, we could now see just what we needed to focus on better. And just as important, our students began to experience the conceptual change that happens when they realize that their preconceptions no longer make sense to them and they start to construct new explanations.

Now we use the probes to try to improve the quality of our district assessments, and we look for evidence that students understand key ideas in science that may have been riddled with misconceptions in the past. We match our summative data to our formative data and identify patterns and discrepancies. *Accountability* isn't such a scary word anymore. We now have the right tools and processes to take the guesswork out of assessment and ensure that our students are ready to "show what they know."

As for the future, although we realize that not all of our students are going to leave our district with plans to be scientists, the chances are now much better that they will leave being science literate, ready to use their knowledge and skills to understand real-world issues and problems that require an understanding of the basic principles of science.

The above vignette is a composite account, drawn from the many stories we have heard from science educators who are using the *Uncovering Student Ideas in Science* series. It shows how formative assessment can provide valuable information to teachers and students to promote learning and inform instruction, creating a more balanced system of assessment that does not overly rely on summative assessment. Assessment isn't only about testing, and this book is not about assessments that are graded and then used to pass judgment on students about the extent to which they have achieved a learning goal. This book is about using a type of assessment, called a *probe*, for diagnostic and formative purposes.

The probes are selected-response items specifically designed to reveal student misconceptions that the research literature on student learning has documented. Throughout this book and the other three books in the series, we use the word *misconception* in a general way to refer to the ideas students bring to their learning that are not yet fully formed and not scientifically correct at the level we would expect. Other words to describe students' ideas include *preconceptions, naive ideas, partially formed ideas, facets of understanding,* and *alternative conceptions*. Although teachers tend to use the word *misconception* in a pejorative way to describe students' ideas that are not the same as the scientific ideas we want them to understand, misconceptions can be useful if we use them to build a bridge between where students are in their thinking and where we eventually want them to be.

Introduction

The assessment probes in this book are designed to help teachers build that bridge. The bridge begins with finding out what students think about important ideas in science that they will use throughout their learning. Although the probes were written primarily to target a K–12 student audience, they can be used with adult learners as well, including university students and science teachers participating in professional development. The previous three volumes provided background information for teachers on formative assessment. We encourage you to collect all the books in the series and read the introductory material to expand your understanding of assessment and its role in teaching, learning, and professional development.

In addition to the 25 probes in each volume, Volume 1 contains an overview of formative assessment—what it is and how it is used. Volume 2 introduces ways to integrate assessment with instruction, and Volume 3 provides an introduction to using the probes, including in professional learning communities. Because we often get the question "But what about testing?"—by which teachers generally mean summative assessment—in the introduction we address the link between formative and summative assessment.

Summative assessment is a pervasive topic that includes everything from statewide accountability tests, to local assessments and district benchmark tests, to everyday classroom tests. To grapple with what seems to be an overuse of graded quizzes and testing, educators need to change their view of assessment to one that is about information. The more information we have about students, from both summative and formative assessments, the clearer the picture we have about student learning.

It is important to remember that these probes are used to elicit students' ideas, engage students in discussion about their ideas, and monitor how students' conceptions are changing throughout instruction. They are not intended to be graded. Once you pass judgment on the student with a grade, research shows that the student's learning often shuts down (Black et al. 2003). Use these probes to gather information about your students and motivate them to open up and share their ideas. Grading a formative assessment often does the opposite. When low achievers get back their papers with low scores, the message is, "You are not good enough," and their desire to learn fades. Likewise a good grade also shuts down students' thinking. As long as students see a passing grade, they often ignore the teachers' comments that may indicate ways to improve their work or challenge them to think further. If we are going to use formative assessment effectively, and distinguish it from summative assessment, it is important to get over the pervasive habit of grading every piece of work. Although there are times when it is important to grade work, the probes are not intended for that purpose.

Assessments fall into three different types: formative, summative, and diagnostic. When data are used by teachers to make decisions about next steps for a student or group of students, to plan instruction, and to improve their own practice, they help inform as well as form

practice; this is formative assessment. When data are collected at certain planned intervals, and are used to show what students have achieved to date, they provide a summary of progress and are summative assessment (Carlson, Humphrey, and Reinhardt 2003, p. 4).

But what about the third type, diagnostic assessment? Diagnostic assessment is used to uncover a misconception or learning difficulty. When used just for this purpose, it is diagnostic. However, diagnostic assessment becomes formative when the information revealed by the assessment is applied to a situation. Another way to look at the three types and their purposes is to use an analogy.

Have you ever watched the Fox television medical drama *House*? Dr. House is a curmudgeonly medical genius who works with an accomplished group of medical diagnosticians who assess a variety of mysterious illnesses in a teaching hospital. What Dr. House does is very similar to what a teacher does when using the probes. Dr. House and his team collect a variety of data to diagnose an illness a patient has. They often use medical probes to look into the body and see things they would not ordinarily be able to see. The data reveal to Dr. House and his team the cause of the illness. Similarly, the teachers' use of the probes uncovers a variety of problems that may not be obvious to other practitioners who do not have the right tools or deep understandings about student learning.

Naturally, Dr. House and his team do not stop with the diagnosis. They want the patient to get better, so they prescribe the best course of treatment that will take care of the medical problem. The treatment is informed by the diagnosis and any additional data on the patient's condition. Dr. House and his team closely monitor the patient for improvement. In educational assessment, this would be the formative assessment—moving beyond the diagnosis to inform the instructional strategies and direction a teacher will use to help students with their learning problems and achieve conceptual understanding.

At the end of a course of treatment, the patient often comes back for a follow-up checkup to see if he or she is cured. This usually happens after the treatment has ended. In education, summative assessment is usually given after a sequence of instruction or at the end of a course to find out how well a student has learned.

Of course, science assessment is not a life or death decision-making process, and today's teachers would not communicate with their students in the way Dr. House abruptly confronts his patients. Teachers' assessments, however, do involve carefully made choices that have a potentially huge impact on student learning. Assessment is not about just testing anymore, and practicing for test taking may not lead to constructive changes in the classroom. Relying on external test results doesn't help much to improve learning in the immediate sense. Traditionally, summative assessments have not helped teachers adjust instruction for individual students because it takes too long for the data to be returned to schools. Formative assessment, such as the probes in this book, can be used on a regular basis to monitor student progress and modify instruction when it is needed.

Teachers and schools are now looking at classroom assessment more closely and at how to best improve teaching and learning so that the assessments we give at the end of a unit or at the end of a year are better matched to what students have been learning, not what they memorized for a test. By linking diagnostic and formative assessments, such as the probes in this book, to the standards being assessed on summative assessments, teachers get a better picture of what they need to do to move students toward the development of conceptual understanding so that students can be successful on summative assessments.

As depicted in the vignette at the beginning of this introduction, some schools are beginning to shift away from a rigid accountability system and move toward the use of more assessments, such as the ones in this book, that look at the whole picture of student learning. When teachers and districts begin to align these formative assessments with the standards assessed on summative assessments, the student data from these probes can provide powerful predictors of readiness for summative assessment. When teachers design their summative assessments with the standards in mind and a picture of what success is, they can use the assessment probes prior to and throughout instruction to monitor how well students are moving toward the learning targets—and adjust their teaching accordingly. As students become more metacognitive about what they know, think they know, and do not know, they take more responsibility for their learning, which eventually improves their performance on measures of achievement.

As a noted superintendent once succinctly said, "Schools are data rich and information poor" (DuFour, Dufour, and Eaker 2005 p. 40). We hope that the probes in this book will uncover information about your students' thinking that will provide a gold mine of useful data for your classroom and for the improvement of your school's science program. Forging a stronger link between formative and summative assessment will create a better assessment balance and lead to your desired results. Are the desired results improved test scores? They may be the "trees" you are looking at, but the real aim is to look through the trees to see the whole forest. In that forest is the "big picture" of learning that will produce science-literate adults. As you use these probes, remember to look beyond the trees and into the forest!

References

Black, P., C. Harrison, C. Lee, B. Marshall, and D. Wiliam. 2003. *Assessment for learning.* Berkshire, UK: Open University Press.

Carlson, M., G. Humphrey, and K. Reinhardt. 2003. *Weaving science inquiry and continuous assessment.* Thousand Oaks, CA: Corwin Press.

DuFour, R., R. Eaker, and R. Dufour, eds. 2005. *On common ground: The power of professional learning communities.* Bloomington, IN: National Education Service.

Mundry, S., P. Keeley, and C. Landel. 2009. *A leader's guide to science curriculum topic study: Designs, tools, and resources for professional learning.* Thousand Oaks, CA: Corwin Press.

Wiliam, D. 2008. Improving learning in science with formative assessment. In *Assessing science learning,* eds. J. Coffey, R. Douglas, and C. Stearns. Arlington, VA: NSTA Press.