

**EDUCATIONAL LEADERSHIP**[Main](#)[Current Issue](#)[Archives](#)[Upcoming Themes](#)[Write for EL](#)[Contact](#)[Subscribe](#)March 2014 | Volume **71** | Number **6****Using Assessments Thoughtfully** Pages 20-26[Issue Table of Contents](#) | [Read Article Abstract](#)

Thoughtful Assessment with the Learner in Mind

Jan Chappuis

Assessing thoughtfully means keeping a continual focus on each learner's needs.

Many of us didn't begin our education careers with expertise in classroom assessment. Our preservice preparation focused primarily on the act of instructing, defined as planning and delivering lessons. Assessments typically came from the textbook materials, they were administered after instruction was complete, and their results took the form of a grade.

Although we've developed a more robust understanding of classroom assessment options, we can still be caught off guard by misapplications of well-intended practices. The emphasis on rigor, data-based decision making, and paced instruction may lead teachers to make serious errors in assessing student learning. Here's how to address these issues thoughtfully.

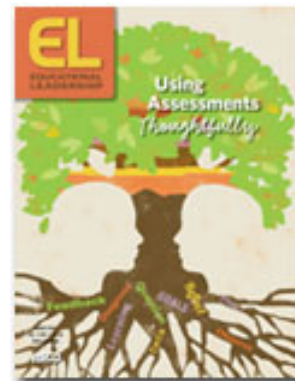
Issue 1. Ensuring Rigor

Over the years, we've repeatedly heard that we need to ratchet up the rigor in our content standards. The Common Core State Standards and the new Next Generation Science Standards have been developed with increased rigor as a primary goal. This has led to a demand for equally rigorous assessments.

Take, for example, Grade 6 Writing Standard 8:

Gather relevant information from multiple print and digital sources; assess the credibility of each source; and quote or paraphrase the data and conclusions of others while avoiding plagiarism and providing basic bibliographic information for sources. (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p. 44)

Faced with a complex standard such as this one, how might I adapt a previously developed research project related to sources of renewable energy? The original assignment asked students to choose one source of

[BUY THIS ISSUE](#)[Share](#) |

renewable energy, learn more about it, identify its advantages and drawbacks for large-scale use, and share their learning with their classmates through a report. In adapting the project to fit this writing standard, I might think it's enough to require students to use at least four credible sources drawn from both print and digital media, to include information in both paraphrased and quoted form, and to prepare a bibliography that includes all the sources they used. Then I might plan instruction in each of the following:

How to look for sources of information.

How to paraphrase.

How to properly punctuate quotations.

How to prepare a bibliography.

However, if I don't look carefully at the types of thinking required by the standard, I most likely will miss teaching and assessing at the appropriate level of rigor. If I examine this writing standard more closely, I find four separate learning targets. The first learning target—*gather relevant information from multiple print and digital sources*—requires that students engage in evaluative reasoning to judge relevance. Without a closer look, I may have taught about types of print and digital sources, rather than about the concept of relevance. The second learning target—*assess the credibility of each source*—also requires evaluative reasoning, this time to evaluate credibility. Without a closer look, I may have simply required that the sources be credible. The third learning target—*quote or paraphrase the data and conclusions of others while avoiding plagiarism*—requires that students know how and when to quote or paraphrase information. Without a closer look, I may have taught the procedural knowledge without teaching the thinking that underlies the effective use of quotes and paraphrases.

So when we take the content standard apart and classify each part according to the type of learning it calls for, we can more clearly see what we need to teach. Do my students understand the concept of relevance? What guidelines can I provide to help them evaluate potential sources? What practice with determining relevance should students do before beginning the project? Do they understand what to look for to determine credibility? What guidelines can I offer?

If we have standards calling for deeper learning, it stands to reason that we'll need assessments that do so, too. But if we believe we can address the issue of rigor by simply giving rigorous assessments, we've missed the point, which is to help students master that deeper learning.

Rigor resides in the standards.

Rigor in assessment must be preceded by instruction in the type of thinking called for by the rigor of the standards.

→ *Assessing thoughtfully means making sure we teach students how to accomplish the rigor embedded in our standards before we hold them accountable for having mastered that rigor.*

Issue 2. Using Data Wisely

Gathering Diagnostic Assessment Data

Research over the last decade (Hattie, 2009) has shown that gathering evidence of student learning during instruction can lead to improved achievement. Although many assessments are administered with the intent of generating diagnostic information, not all are capable of doing so (William, 2013).

For example, consider the following test item:

Which fraction is largest?

1. $\frac{1}{3}$
2. $\frac{2}{5}$
3. $\frac{7}{9}$

Many 4th graders would most likely be able to choose the correct answer *c* because both the numerator and the denominator are the largest numbers in the set. However, this set of answer choices doesn't accurately differentiate between students who understand the concept and students who don't. Students could get it right for the right reason (because they understand that the *relationship* between the numerator and the denominator determines size) or for the wrong reason (because they believe that size is determined by the numerator or denominator). The problem doesn't help ferret out misconceptions that may be lurking.

If we plan to use information from our assessments, the information must first be accurate. When students can get an item right for the wrong reasons, we haven't examined the answer choices carefully enough. And if the wrong answer choices don't give us information about what problems further instruction should address, the item doesn't have any diagnostic power. Assessments that yield actionable diagnostic information provide results that identify specific learning needs. Such assessments are *instructionally tractable* (Andrade, 2013).

Consider the answer choices in this problem:

Which fraction is largest?

1. $\frac{2}{1}$
2. $\frac{3}{8}$
3. $\frac{4}{3}$

Students who understand that the relationship between the numerator and the denominator determines size will choose answer *a*. Students who use the denominator to determine size will likely choose answer *b*. Students who use the numerator to determine size will likely choose answer *c*. With answer choices like these, you know who does and doesn't understand magnitude in fractions, and you also know what to focus on with students who've selected either of the two wrong answers.

The diagnostic capability of an assessment is key to its instructional traction. Assessments of all types—selected response, written response, performance assessment, and personal communication—can be designed to provide results that are instructionally tractable, both as a formal assessment and as an informal lesson-embedded activity. (See fig. 1 on p. 25 for examples of multiple-choice item formulas designed to identify specific flaws in reasoning.)

Figure 1. How to Create Test Items with Instructional Traction

INFER
Question: Which idea could you infer from the text? or, Which idea does this selection suggest?
<p>Possible Answers:</p> <p>The right answer: A guess based on evidence found in the passage.</p>

Distractor: A guess that includes a word or concept copied directly from the passage but that is not supported by the meaning of the passage.

Distractor: A guess that might seem reasonable but for which there is no evidence in the passage.

SUMMARIZE

Question: Which sentence best summarizes what this passage is about?

Possible Answers:

The right answer: A statement of the main ideas of the passage.

Distractor: A statement of one main idea, not sufficiently broad to represent the whole passage.

Distractor: A statement including an idea not found in the passage.

Distractor: A statement of one fact or detail from the passage.

IDENTIFY CAUSE AND EFFECT

Question: Which sentence best explains why _____(an event or action) happened?

Possible Answers:

The right answer: A plausible statement of causation based on evidence from the text.

Distractor: A statement of causation that is not supported by evidence from the text.

Distractor: A statement that offers another effect rather than a cause.

*Source: From *Seven Strategies of Assessment for Learning* (2nd ed., in press), by Jan Chappuis, 2015, Upper Saddle River, NJ: Pearson Education. Copyright 2015 by Pearson. Reprinted with permission.*

→ *Assessing thoughtfully means requiring that assessments we intend to use diagnostically have sufficient instructional traction.*

Using Diagnostic Information to Plan Instruction

Students encounter difficulty during instruction for a variety of reasons. Which actions teachers take depend on the type of learning needs our diagnostic information has revealed. It's not enough to know that students haven't mastered something yet—it's also important to know the type of problem that is standing in the way of mastery so that we can plan appropriate next-step instruction.

We can think of the errors that students make as falling into three categories: errors due to incomplete understanding, errors due to flaws in reasoning, and errors due to misconceptions (Chappuis, 2014).

Errors due to incomplete understanding. These kinds of errors are fairly easy to recognize and correct. They are not so much errors as they are lack of knowledge: The student's work shows at least partial knowledge, which provides an opportunity for further instruction.

For example, when primary children start stringing written words together into complete thoughts, they might put a period between each word instead of at the end. They understand something about using

periods, but they haven't learned their proper use. Teachers shouldn't mark the periods as wrong, but, instead, should move students toward stringing words together as units of thought and then teach them to use periods where a unit of thought ends.

Likewise, elementary students studying ecosystems may not know that river and stream water is called *fresh* water—they may call it something like *plain* water. Teachers can simply supply the new term and help students understand it by distinguishing it from salt water.

For an error that is caused by incomplete understanding or lack of knowledge not yet taught, rather than labelling it as an error, teachers can introduce the new information immediately or, as in the case with the primary students and period use, when it becomes developmentally appropriate.

Errors due to flaws in reasoning. Addressing errors that are the result of flaws in reasoning requires thinking carefully about the type of reasoning involved and then helping students recognize typical errors for that particular pattern of reasoning.

For example, when students are asked to generalize, they often either overgeneralize or don't generalize at all. When they're asked to infer, they often don't read carefully to look for evidence that will support an inference. When they're asked to summarize, they often leave out important points or include unimportant ones. To overcome flaws in reasoning, teachers should help students understand the salient features of the pattern of reasoning and let them examine examples of the flaws so they can more easily recognize them before the students begin practicing with that pattern of reasoning in the context of a given subject.

In the case of generalization, a teacher might signal reasoning flaws by making the statement, "All dogs have four legs." But some dogs—due to injury or illness—have fewer legs, and a few students most likely would point this out. By asking students to come up with a statement they could agree on—such as, "*Most* dogs have four legs"—the teacher can lead students to conclude that her broader claim is an overgeneralization, a claim that goes too far.

To help students distinguish between overgeneralizations and appropriate ones, the teacher might assign students a short passage to read that's accompanied by three statements, two of which are overgeneralizations and one of which is an appropriate generalization. After reading the passage, students could work in pairs to determine which is which and to explain their choices.

To overcome the tendency to draw an inference based on too little information, a teacher might use an everyday example, such as making a guess about a student's favorite color on the basis of one article of clothing that student has chosen to wear and asking students to identify why that guess might not yield a defensible inference. Students could then work with a partner to examine a set of possible inferences drawn from a short reading passage to determine which are based on sufficient evidence and which are not.

When students' summaries leave out important points or include unimportant details, the teacher might create a summary of a familiar story, such as the story of Goldilocks and the Three Bears, that has one of those problems. The teacher would explain to students that a good summary is a brief description of the main ideas and then ask students to determine whether the summary of the Goldilocks story is a good one—and if it's not, why not.

With errors due to flaws in reasoning, give students time to analyze examples of typical flaws as well as examples of good reasoning before asking them to practice that type of reasoning themselves.

Errors due to misconceptions. Misconceptions involve students either having learned something inaccurately or having internalized an explanation for a phenomenon that doesn't fit with current best thinking. Basically, with a misconception, students have learned or internalized something that they believe to be correct but that isn't.

The challenge with misconceptions is to correctly identify them and then plan lessons to dislodge them. Misconceptions are stubborn: They can't be corrected by papering over them. To illustrate, let's look at a

misconception that's common in middle school science. Newton's first law of motion states that a force is not needed to keep an object in motion, yet many students (and adults) will tell you that if an object is in motion, it will require a force to stay in motion, which seems like common sense. (Aristotle thought this, by the way.) Memorizing the principles of the first law—"an object at rest will stay at rest" and "an object will continue with constant velocity unless acted on by an unbalanced force"—is generally not enough to counter what our senses tell us about force and motion: If you want a book to keep moving across a table, you have to keep pushing it.

One effective approach to dislodging misconceptions (Hattie, 2009) is to first create an awareness of the misconception by providing students with an experience—such as a demonstration or a reading passage—that runs counter to the misconception in some way. The teacher might then ask students, Where does the experience contradict what you think is right? to identify the misconception and contrast it with the correct interpretation. Finally, when students are able to do so, have them explain why the misconception is incorrect.

Misconceptions, whether in science, social studies, mathematics, language arts, or any other discipline, require an intentional approach tailored to the nature of the misconception because the teaching challenge is to cause conceptual change—to have students *give up* the inaccurate conception they currently hold in favor of an accurate one.

→ *Assessing thoughtfully means searching out what students' work tells us about their learning needs and planning further instruction accordingly.*

Using Common Assessments as Diagnostic Information

Over the past decade, many schools and districts have engaged in various forms of data-driven decision making in which they gather evidence of student achievement and then discuss next steps. Teachers often administer a common formative assessment designed to cover a set number of content standards so they can meet to discuss the results.

One typical problem here is that the assessment may or may not measure what the teacher has taught. Formative assessment information that teachers gather during learning must reflect the content standard that students are in the midst of learning. Teachers need to be able to gather this information informally every day and formally on a regular basis. If teachers give common formative assessments when they haven't taught all the content standards represented on that assessment or aren't at a point in instruction where the information is actionable, they've put accountability before the thoughtful use of formative assessment. When teachers are required to give predetermined common assessments at a predetermined time, accountability for covering material has superseded instructional use.

We can overcome this problem by identifying which learning target each item on a common assessment measures and by checking to be sure whether formative information about that learning target is needed before administering that part of the common assessment.

→ *Assessing thoughtfully means using common formative assessments only when they are instructionally relevant both in content and in timing.*

Issue 3. Keeping Learning Moving Forward

Pacing

After we've planned and delivered the lesson and after the students have done something in response, we know whether they've mastered the intended learning. If the evidence suggests that they haven't, our options

are to grade the assessment and move on, reteach the lesson, or figure out what the students' current learning needs are in relation to the content standard and teach to those needs. Research consistently favors option three (Hattie, 2009; Wiliam, 2013).

However, when the content and rate of instruction are controlled by a pacing guide, there's little opportunity to do anything but grade the work and move on. Many pacing guides have been developed with an incomplete understanding of the role that diagnostic assessment information plays in effective teaching and learning. If teachers are to comply with a district, building, or department mandate to adhere to the established pacing guide, they simply have no time to slow down instruction to the pace of learning.

Because it's often not a straight shot from instruction to mastery, good teaching includes knowing how to keep learning moving forward in the face of difficulty. John Hattie (2009) calls this part of teaching a *feedback loop*. An effective feedback system begins with feedback *to* the teacher *from* the students about what they have and have not learned. Feedback to the student may be the appropriate next step, but when student work doesn't demonstrate at least partial mastery, feedback is not the best response. Hattie identifies the teacher's willingness to seek *disconfirming evidence*—to actively look for evidence of those parts of the learning that need additional focus—as the most crucial factor in helping each student learn.

This should be the main purpose of our diagnostic assessments, whether they're formal events, such as quizzes and tests, or informal events, such as questioning, dialogue, homework, and quick-checks.

→ *Assessing thoughtfully means ensuring that our pacing guides accommodate the teaching and learning needs that our diagnostic assessments have identified.*

About Grading Too Soon

Students make decisions every day about who they are as learners, whether to try, and whether they're good at specific subjects or school in general on the basis of what we do in response to their assignments and assessments.

Let's assume that students don't already know the things we're teaching. It's reasonable to assume that they'll need instruction followed by practice, which they'll often not execute perfectly at the start. If students receive a grade on their practice work, those who don't do well will tend to conclude that they aren't very good at the task or subject. And if we average the grades on earlier attempts with later evidence showing increased mastery, we won't accurately represent students' true level of mastery. As a result, some students may simply give up trying when they realize they can't overcome the damage done to their grade by their initial lack of understanding.

Yet, that's the premise we began with: They aren't good at it ... yet. Grading too soon in the learning process can lead students to the damaging inference that being good means not having to try and that if you have to try, you aren't smart in the subject. If one of our goals is to get students to try, then trying shouldn't result in the punishment of a low grade assigned too soon.

During the learning, students' work will reveal some combination of incomplete understanding, flaws in reasoning, or misconceptions. Our job as teachers is to examine student work and offer sufficient penalty-free practice time, reteach and redirect students when needed, and provide both success and intervention feedback as called for.

→ *Assessing thoughtfully means giving students sufficient time and opportunity to practice and improve through further instruction and feedback before holding them accountable for having mastered the learning target.*

From Issue to Opportunity

The heightened rigor of the new content standards, using assessment information for data-driven decision making, and using pacing guides to map instruction can all contribute to better learning for our students—if we pay attention to what the learner needs. Assessment, when carried out thoughtfully, is at the heart of teaching well to meet those needs.

References

Andrade, H. L. (2013). Classroom assessment in the context of learning theory and research. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 17–34). Los Angeles: SAGE.

Chappuis, J. (2014). *Seven strategies of assessment for learning* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects. Washington, DC: Author.

William, D. (2013). Feedback and instructional correctives. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 197–214). Los Angeles: SAGE.

Jan Chappuis is an independent consultant specializing in classroom assessment practices and author of *Seven Strategies of Assessment for Learning*, 2nd ed. (Pearson, 2014).

KEYWORDS

Click on keywords to see similar products:

[assessment](#), [learning](#)

Copyright © 2014 by ASCD

Requesting Permission

- For **photocopy**, **electronic and online access**, and **republishing requests**, go to the [Copyright Clearance Center](#). Enter the periodical title within the "**Get Permission**" search field.
- To **translate** this article, contact permissions@ascd.org