EL Action Pack on Assessment

Featuring three top-rated articles from Educational Leadership Magazine

Preventing Feedback Fizzle........................................2
Thoughtful Assessment with the Learner in Mind ..........10
How I Broke My Rule and Learned to Give Retests......19

Educational Leadership (EL) is ASCD’s flagship publication. With a circulation of 130,000, EL is acknowledged throughout the world as an authoritative source of information about teaching and learning, new ideas and practices relevant to educators, and the latest trends and issues affecting prekindergarten through higher education.

Subscribe today!
PREVENTING FEEDBACK
FIZZLE
There's more to feedback than just crafting thoughtful comments. Here's how to avoid common pitfalls and make the most of feedback.

Feedback is certainly about saying or writing helpful, learning-focused comments. But that's only part of it. What happens beforehand? What happens afterward?

Feedback that's helpful and learning-focused fits into a context. Before a teacher gives feedback, students need to know the learning target so they have a purpose for using the feedback they receive. Say you're trying to teach students how to identify the main idea in expository text. If a student isn't trying to learn how to do this with the text he or she is working on, your feedback about emphasizing a certain point (such as, “Tell us more about the Articles of Confederation”) will seem like something you want the student to do to comply with your wishes, instead of something the student needs to learn (such as why the Articles of Confederation are so important to the main idea of the text).

But there's another essential component to effective feedback. After receiving feedback, students need the opportunity to digest, understand, and use it.

It Starts with a Target
Before feedback occurs, students need to know what they're trying to learn. Learning targets are student-friendly descriptions—through words, pictures, actions, or some combination of these—of what you intend students to learn or accomplish in a lesson. They're connected to a performance of understanding—something the student actually does to pursue the target—as well as to accompanying criteria for good work that students use to gauge their progress toward the goal. Learning targets are not instructional objectives, which teachers use for instructional planning and which can span one or more lessons.

For example, suppose a teacher's goal is, “Students will recognize and draw shapes having specified attributes, such as a given number of angles or a given number of equal faces” (part of Common Core Mathematics Standard 2.G.1). Students will work on this goal for a whole unit. There will be lessons (and objectives in teacher language) about identifying angles, faces, and other attributes of various figures; lessons about congruence; lessons about solving problems; and so on.

One day, students are learning that congruent figures have exactly the same attributes. Here's how the teacher might state the learning target: “I can draw a figure that's exactly the same size and shape as an example.” (Notice that the learning target is expressed from the student's—and not the teacher's—point of view.) The teacher shows students their learning target using correct and incorrect examples of congruent figures. Then she engages students in a performance of understanding. Using grid paper, students are to exactly reproduce two irregular polygons and receive feedback from a partner about whether their figures are the same as the examples. Then each student must create an original irregular polygon on grid paper for his or her partner to replicate.
Students turn in their final work with explanations for why their figures are exactly the same as the examples provided.

Notice that the teacher told the students what the learning target was (using the I-can statement) and showed the students what the learning target was (using the correct and incorrect examples). Then the students had the opportunity to show themselves what the learning target was and how they were doing (the performance of understanding, the grid paper exercise). The criterion, in this case, was built in (Are the two figures exactly the same?).

If students do an assignment simply because you asked them to, that’s compliance. Compliance is reactive, not proactive. Of course, students should do what you ask, but they won’t learn much unless they understand why you’re asking. When you say, “You can show how well you understand what a food chain is by drawing one and then solving some what-if problems related to it. Here’s an example,” students aren’t just complying, but also learning about food chains and producing evidence of what they’re learning.

Feedback can’t work if students aren’t trying to reach a learning target—or don’t know what the target is, or don’t care. In that instance, information is an answer to a question students aren’t asking. Feedback without a learning target is just somebody telling you what to do.

**When Feedback Fizzles**

When the learning target and the performance of understanding don’t match exactly and the criteria aren’t clear, students often experience feedback as evaluation or grading rather than information for improvement, as in the following example.

A middle school mathematics teacher was teaching his pre-algebra students how to solve one-step equations with one variable. The concept he wanted them to learn was that using inverse operations will isolate a variable on one side of the equation and lead to an efficient solution. The class did several examples together, and then the teacher had the students do a problem set individually. The directions on the problem set read, “Solve. Show all steps.” There was no mention of inverse operations except as implied in the term steps. In an effort to keep calculations easy, many of the problems could be solved with mental arithmetic.

One of the problems read \( m + 8 = 15 \). Quickly calculating this in her head, one student wrote \( m = 7 \). The teacher marked the problem wrong. The student’s first reaction to this feedback was, “That’s not wrong!”

The fact is, they were looking for different things. The teacher was looking for evidence of the use of inverse operations; he wanted the student to understand how subtracting 8 from both sides of the equation would solve the problem—and he wanted the student to show this work. The student was looking for an efficient solution to the problem, which she achieved. The teacher’s feedback wasn’t descriptive; it was an evaluation.

It’s not that the teacher’s objective was wrong. My point is that the teacher didn’t communicate the objective as a learning target the student should aim for, and the result was ineffective feedback. The student got angry instead of looking to learn more.
When Feedback Sparkles

When the learning target and the performance of understanding match exactly and the criteria are clear to both students and teacher, teachers can give feedback that students understand and use right away, as in this example.

Erica Smith teaches a Title I, extended-day kindergarten class. Her learning target for students for one day's lesson was this: “I can recognize whether there's a short â sound in the middle of a word.” For this lesson, the teacher described and modeled the criteria for students, saying, “We’ll know that the word has the â sound in the middle if our mouth is open wide and our tongue is flat when we say the middle sound.” She modeled the strategy of “stretching” the word with her hands (fl – a – t) to more clearly emphasize the middle sound.

The performance of understanding was for students to stretch words and decide whether there was an â sound in the middle. They began by pronouncing some words as a group and shaking heads (yes/no) to identify whether the middle sound was â. Then students did this several times individually, with Erica giving oral feedback. The final product was completing a written assignment that had five picture words for students to indicate whether there was an â sound in the middle.

One student, Marisa, made a mistake with the word flag; she said there was no â sound in the middle. Erica's feedback began with the question, “Tell me the picture name.” Marisa responded “flag,” showing Erica that the source of error wasn’t that Marisa didn’t know the picture name. Next, Erica asked Marisa to stretch the word flag. When Marisa did, she realized it did have an â sound in the middle. When Erica asked, “How do you know?” Marisa modeled an open mouth and a flat tongue, showing she used the criteria herself. Using a series of questions to delve into a child's understanding—what Clarke calls “scaffolded prompts”1—is an effective way to extend students' learning.

It Finishes with Use

Feedback can't be left hanging; it can't work if students don’t have an immediate opportunity to use it. In my experience, teachers are better at giving immediate feedback than at setting up opportunities for students to use it.

One exception to this is teachers who use the writing process. These teachers already know the “immediate opportunity to use” principle. Students regularly incorporate first-draft feedback into revisions for their final copies.

This approach works in a wide variety of situations, however, not just in writing class. Whether students are writing reports or doing projects, the teacher should give them feedback on drafts and partial products so they can incorporate the feedback into their final products, revise them, and then reflect on how the changes improved their work.

When Feedback Fizzes

When students get feedback on a performance that’s not followed by an opportunity to demonstrate the same knowledge or skills, feedback will fail. Feedback “so they know better next time” is a waste of energy. This isn’t the students’ fault, and it doesn’t mean they didn’t take your feedback seriously. It’s just a characteristic of how people learn.2
For example, a middle school reading and language arts teacher wanted her students to learn how to summarize nonfiction text (her instructional objective). She told her students that “summarizing nonfiction text” was their learning target, and she gave them a worksheet that divided a chapter in their social studies text into five sections (for example, “Summary of pages 321-324,” “Summary of pages 325-337,” and so on), with blank spaces under each for students to write their summary. She reminded the students that a summary restates the big ideas of the text, eliminating details. She told them they would know they had succeeded when they could write their own summaries of chapter sections, using those criteria (big ideas, no details), and get a minimum grade of 75 percent.

This example is a double fizzle: To begin with, the teacher never provided a clear, shared learning target and criteria. Then, to compound things, the feedback came as a grade at the end of the learning episode. Because summarizing textbook information is a basic skill, the teacher reasoned, the students would use the feedback they received in some as-yet-unspecified future textbook reading.

First, consider the learning target. “Summarizing nonfiction text” isn’t a daily learning target; it’s a major skill that develops over the course of a student’s education. Moreover, the students were given no examples or models, just told that a summary contains the big ideas from a text. The learning target should probably have been something like this, expressed from the student’s point of view: “I can summarize information on ecosystems from my textbook, and I’ll know I can do it when I can put all the important ideas in one section of the textbook in a single paragraph.”

Second, consider the performance of understanding, what the students were actually supposed to do to move toward their target and show evidence of having learned it. It was just a list of five page ranges, the supposition being that when students read text, they’ll be able to capture main ideas.

Third, consider the criteria. Using big ideas and eliminating detail are descriptions of quality summaries of the sort the teacher envisioned, but “getting a minimum of 75 percent” is an evaluative criterion that is of no help to students as they’re writing their summaries.

Finally, consider the summative or end-of-story nature of the feedback. This is what breaks my heart about this example. What the teacher actually wrote on her students’ scored worksheets was very thoughtful, descriptive feedback, with suggestions for next steps. For example, on one set of summaries that she awarded a grade of 3 out of 4 she wrote, “I can see you made an effort to keep your summary brief, and that was a goal of this lesson. If you had told us how the Everglades was formed and then almost destroyed, this would have given you a 4.” Just looking at this feedback, without knowing the rest of the story, you might judge it effective.

But there was no next step. The assignment was done, the students were finished, and the feedback was moot. If the teacher had given the same feedback as an intermediate step, before the final set of summaries was due, the student could have used it to revise the work before turning it in for grading. An alternative, and probably a better use of time, would have been for the teacher to ask students to write a summary of one of the five sections she listed and turn it in for feedback. Students could then have used her feedback to revise that one summary and, with this knowledge under their belts, write the additional four summaries.
Good Feedback Is ...

1. *Timely.* It arrives while the student is still thinking about the work and while there’s still time for improvement.

2. *Descriptive of the work,* not the student personally. It focuses on one or more strengths of the work and provides at least one suggestion for a next step. Don’t assume that your students know what they did well and that they only need corrections or fixes.

3. *Positive.* It shows how learning is a journey forward, and it’s honest about both strengths to build on and weaknesses to improve. Its tone conveys to the student that the teacher thinks of him or her as an active learner.

4. *Clear and specific.* It’s specific enough so the student knows what to do next, but it leaves the student with some thinking to do.

5. *Differentiated.* It meets the needs of each student with respect to the current work. For some students, a reminder is all that’s needed for a next step; others may need prompts or examples.


---

**When Feedback Sparkles**

Teachers set up feedback to be effective when the learning target and the performance of understanding match, when students have a clear idea of the criteria for their performance and get immediate feedback on that performance, and when they have an opportunity to perform this skill or activity again. Here’s an example.

A 9th grade physical science teacher wanted her students to learn how varying conditions affect projectile motion. This was her objective. Her learning target for students was that they would be able to predict how projectiles would move. For a performance of understanding, she asked students to predict the effects of four projectile characteristics (the object’s angle of launch, initial speed, mass, and diameter, both with and without air resistance) on three characteristics of the projectile’s motion (how far it goes, how high it goes, and how long before it hits the ground—or, range, height, and time, respectively). In groups, students wrote a statement about how they predicted each of the four characteristics would affect the three aspects of motion and explained their reasoning.

The next learning target was that students would be able to assess the accuracy of their predictions and reasoning. The performance of understanding was a lab. Students used a web-based simulation in which they changed one input variable at a time and created a table to record their results. Then they compared the predicted and observed results for
each input variable and wrote down whether the simulation results supported their initial reasoning or not. The criteria were accuracy of comparison and soundness of scientific reasoning.

Students used their prediction sheets and data tables to write lab reports, and they submitted rough drafts to the teacher. She gave feedback on the substance of the reports—that is, on students’ observations and reasoning about how changing the characteristics of objects affected their projectile motion. Her feedback was not about lab report format or “correctness” of conclusions, but about the observations and reasoning. The feedback was not “giving away answers” but rather pushing students to learn more. For example, on one student’s report she noted, “A larger diameter should have a shorter range, distance, and hang time than a smaller diameter when air resistance is present. How can you show this?” Students then had an opportunity to revise their lab reports before handing them in for a final grade.

**Avoid the Fizzle**

To avoid feedback fizzle, take the following steps.

First, share the learning target and success criteria for each lesson with your students. Make sure your performance of understanding—what the students actually do during the lesson—is a spot-on match with your learning target. This 3 of 4 accomplishes several good things. By sharing the learning target in the assignment itself—and not just in words—students can envision what they’re supposed to learn by looking at what they’re asked to do. As students do their work, they make progress toward the target. This work produces evidence on which teachers can base effective feedback, which students can use, in turn, to self-regulate their learning.

Second, whether your feedback is oral or written, choose your words carefully. Describe the work’s strengths and give at least one suggestion for a next step that is directly in line with the learning target. Use words that suggest the student is an active learner and will make decisions about how to go forward, not words that suggest a student should use the feedback by complying with a request. For example, you might say, “What were you thinking as a writer when you described the tree?” and not, “Why did you write about the tree?”

Third, follow episodes of feedback with immediate opportunities for students to use their feedback, before you give them a grade. For writing and complex projects, students may use feedback for revisions and redos. However, for solving a mathematics problem, applying punctuation rules, balancing chemical equations, and other application-level tasks, revisions and redos are not appropriate because students have already seen the answers. They need to use feedback to tackle other similar problems. They don’t necessarily have to do another whole page of work; sometimes another problem or two is enough to show themselves and you that they’ve been able to use the feedback and are ready to move on.

Put these feedback tips in place—and watch your students sparkle!
Endnotes

“Preventing Feedback Fizzle” by Susan M. Brookhart is reprinted from Educational Leadership, Volume 70, Number 1 (September 2012).

Susan M. Brookhart is an independent education consultant based in Helena, Montana, and a research associate in the Center for Advancing the Study of Teaching and Learning at Duquesne University. Her latest book, coauthored with Connie M. Moss, is Learning Targets: Helping Students Aim for Understanding in Today’s Lesson (ASCD, 2012).

Copyright © 2012 by ASCD
THOUGHTFUL ASSESSMENT WITH THE LEARNER IN MIND
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 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:

1. How to look for sources of information.
2. How to paraphrase.
3. How to properly punctuate quotations.
4. 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 (Wiliam, 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 deter-
mines 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. 2/1
2. 3/8
3. 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.

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 under-
stand 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


“Thoughtful Assessment with the Learner in Mind” by Jan Chappuis is reprinted from Educational Leadership, Volume 71, Number 6 (March 2014).

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

Copyright © 2014 by ASCD
HOW I BROKE MY RULE AND LEARNED TO GIVE RETESTS
Structured choices for retesting can motivate even the lowest achievers.

Six years ago, a conference on standards-based grading and assessment left me with the distinct and slightly confounding sense that my assessment procedures needed to change. As it turned out, one of the assessment practices I had been most wedded to was one I eventually overturned.

The conference, which my principal at the high school where I then taught urged me to attend, reinforced nagging questions I’d long had about traditional grading practices. During the first sessions, I agreed in whole or part with nearly everything presenters said. For years, I’d encouraged students to make practice tests, so the concept of formative assessment was familiar to me. Like the presenters, I’d faced the conundrum of how to equitably grade the bright student who did little homework but scored high on tests, and I’d felt uneasy with the practice of reducing grades for assignments that were handed in late. Then a speaker prompted me to question one of my most entrenched rules: Never offer retests.

On my flight home, as I reflected on what I’d learned, I realized how much retests would collide with the grading procedures I used as a high school history teacher. I saw two obvious stumbling blocks to offering retests: (1) My courses were very content-heavy, and I had little time to spend on retests; and (2) I didn’t know how to maintain an authentic measure of learning if I allowed everyone to take retests.

Pondering how to make it work, I thought of the model of assessment Rick Stiggins had presented. His model made sense to me, especially the three key questions he said students must know how to answer. At the beginning of a unit, all students should be able to answer the question, Where are we going? After an assessment, they should be able to answer the question, Where am I? and after answering both of these questions, the student should be able to answer, How do I close the gap? I used these questions as touchstones as I transformed my testing policy.

**Smooth Sailing on “Where Are We Going”**?
With a little help from my vice principal at the time, Tom Schimmer, this was a relatively easy question to address. In his previous school, Tom had been using student-friendly unit plans that clearly delineated learning targets—what a student needed to be able to do during each unit. I began using learning targets in my senior history courses. I laid out all unit requirements under one of the following headings:

1. Knowledge Targets: What do I need to know?
2. Reasoning Targets: What can I do with what I know?
3. Skill Targets: What can I demonstrate?
4. Product Targets: What can I make to show my learning?
I presented each target as an “I can” statement, which made it easier for students to understand the target and take ownership of reaching it.

Both my students and I found these unit plans incredibly helpful. In the knowledge targets section, students could find all the “Trivial Pursuit” objectives—definitions, dates, names, and other specific information they needed to know. I explained to students that these knowledge pieces were essential to success in the course; any discussion or activity we did in the unit would require them to know these core facts. For example, one knowledge target for our pre–World War II unit was “I can list four conditions in 1930s Germany that resulted in Hitler gaining power.”

Reasoning targets reflect what students can do with what they know. In my courses, these are often the most interesting targets, as students are required to bring knowledge pieces together to form an argument or make a judgment. Terms such as justify, determine, compare, and evaluate are commonly found in this section. An example of a reasoning target from the pre–World War II unit might be “I can explain to what extent the United States followed a policy of isolationism in the 1930s.”

Skill and product targets are relatively easy to determine. Skill targets focus on what students can do to demonstrate understanding, such as make a speech or complete a hands-on map activity. One student’s skill target for the pre-World War II history unit was “I can research a member of the Jewish community living in 1930s Germany and give a two-minute speech on his or her specific concerns or challenges.” Product targets refer to what students make to show learning; for example, a short written description, PowerPoint presentation, or collage of images that represent the social conditions of the 1930s.

My students reacted very positively to these structures. At the end of each unit, we went over the unit’s plan as a class. (In one case, we discovered that we’d missed covering a knowledge target because a fire drill took us out of class!) Students used the targets as study guides by checking off the “I can” statements and determining what they still needed to learn. As one student noted, “I’m able to discover what I know and don’t know before I take the test.”

**Hitting Bumps at “Where Am I?”**

I thought students could answer this question as they always had—by seeing their graded tests and my feedback. The new element would be that students could close the gap by further study followed by a retest. I believed I could administer retests using my existing test structure and rely on my comments to guide students toward improvement. It turned out to be more complicated.

With some trepidation, on returning a set of tests, I announced to my History 12 students that students who were unhappy with their results could see me after class to schedule a retest. Allie was one student who requested a retest, and we arranged a lunch meeting for the following day.

Before her appointment, I looked over Allie’s test, and I knew I had a problem. The test consisted of a few sections. Allie requested to focus only on the first section of 40 multiple-choice questions, in which she had 12 errors. Given the complexity of the question format and the integration of different learning objectives into different sections, I couldn’t
ascertain specifically where her weaknesses lay. Consequently, I couldn’t determine an
efficient and accurate way to retest Allie. Even if I asked a complete second set of random
questions and Allie rewrote the whole test, I still couldn’t guarantee that her second assess-
ment would be a clear replacement of the first.

I ended up interviewing Allie on the individual questions she had missed, trying to see
whether she now understood them better. It was a painful, inefficient process that lasted 30
minutes and didn’t give either of us much insight. I was left with the clear understanding
that I’d better revamp this process.

As a first step, I reorganized my tests. Rather than sticking with my usual formula of separat-
ing each test into sections by type of question—multiple choice, short answer, long an-
swer—I rethought my structure. I settled on separating sections by learning outcomes/major
topics and varying the type of questions within each of these sections. For example, my test
on Franklin Delano Roosevelt (FDR) resulted in the following sections and values:

1. The United States in the 1920s: 11 points
2. Causes of the Depression: 4 points
3. FDR’s efforts to end the Depression: 5 points
4. Reactions to FDR’s New Deal: 7 points
5. The End of the Depression: 6 points

Section 1, for instance, included eight multiple-choice questions and a paragraph worth
3 points that students wrote to a prompt, for a total of 11 points. Section 3 consisted only
of definitions. As I constructed each section, it dawned on me to simultaneously write the
corresponding section to the “sister” test. While I had my head wrapped around causes of
the Depression, for instance, it was easy to make another section on the same topic, also
worth 4 points. By the time I was finished, I had two tests with the same sections and values,
but different questions.

After my students took the first FDR test, I graded it as usual. When I handed it back,
however, the routine changed. I gave each student a tracking sheet (see fig. 1) on which
I had listed the different sections and values of the test. I had students write in the points
they scored on each section and tabulate their percentage score. The last box beside each
section was where students indicated whether they intended to retake that section. Within
a few minutes, students had a graphic representation of their strengths and weaknesses on
each learning outcome. Because students actively tabulated their own section scores, the
classroom atmosphere was a far cry from the disengaged atmosphere so common when
teachers return tests.
### FIGURE 1. Tracking Sheet

**Franklin Delano Roosevelt**  
**Name:** Jon Black  
**Date:** April 3, 2010

<table>
<thead>
<tr>
<th>Topic</th>
<th>Value</th>
<th>Score</th>
<th>%</th>
<th>Retest?</th>
</tr>
</thead>
<tbody>
<tr>
<td>The USA in the 1920s</td>
<td>11</td>
<td>8</td>
<td>73%</td>
<td></td>
</tr>
<tr>
<td>Causes of the Depression</td>
<td>4</td>
<td>3</td>
<td>75%</td>
<td></td>
</tr>
<tr>
<td>FDR’s efforts to end the Depression</td>
<td>5</td>
<td>2</td>
<td>40%</td>
<td>✓</td>
</tr>
<tr>
<td>Reactions to FDR’s New Deal</td>
<td>7</td>
<td>7</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>The End of the Depression</td>
<td>6</td>
<td>3</td>
<td>50%</td>
<td>✓</td>
</tr>
</tbody>
</table>

Total points 23 out of 33. Overall score: 70%.  

**Unit Terms/Preparation**

- [ ] I DID complete all of the terms for this unit on either cards or sheets.
- [✓] I DID NOT complete either the cards or the term list for this unit.
  
  Reason: I didn’t think I needed to; I felt prepared.
- [ ] I DID complete a different form of preparation. Explain: ____________

**Goals and Strategies**

What **overall grade** (percentage or letter) am I hoping to achieve in this course? **85%**

- [ ] I did all that I could to achieve my goal in preparing for this test.
- [✓] I plan to make the following adjustments to increase my grade:
  
  ✓ Complete all vocabulary cards  
  ✓ Make practice quiz to test myself
Progressing Toward Closing the Gap
While I had my students’ attention, I included on each tracking sheet questions about their test preparation, study skills, and goal setting, and then collected the completed sheets. I found that some students admittedly struggled to study effectively. Looking over the sheets, I could determine which students were—and weren’t—using my suggested study routines. If students were not doing assigned homework or not taking time to study, when appropriate I made these actions prerequisites for a retest. I returned a copy of their tracking sheet to all students who requested a retest. Each student went home with a copy of his or her section scores, a list of which sections to study for the retest, and a summary of suggested study routines.

We scheduled students’ retests during class or at lunchtime. Some students selected to retest only one section, others chose to retackle multiple sections, and some left their test score as it was. I offered topic-specific tutorials on areas students missed, at lunch or after school.

At first, struggling learners often chose to retest only one section. I took this opportunity to converse with each struggler about preparation and study techniques and to urge him or her to put in an extra study session independently or schedule a session with me. Because extra studying focused on only one section or topic, the at-risk learner usually perceived it as easier and 3 of 4 shorter and was willing. As low-scoring students began to see dramatic improvement on their retested sections, many displayed heightened levels of confidence and tackled multiple sections on subsequent retests.

This procedure was also a good tool to assess my teaching. If I noticed that most students scored low on a particular section, I took that as a sign that my instruction on that section might need adjustment. As a class, we have revisited and relearned particular sections and I’ve scheduled whole-class retests.

A Few Observations
Since I started revamping my testing procedures, I’ve seen more examples of how the change benefits students and gathered more insights than I could share in one short article. But here are a few of my observations:

1. The ability to retest on specific learning outcomes benefits both low- and high-achieving students. When a struggling learner sees a score of 80-100 percent on one section after a retake, I’ve observed considerable improvements in his or her overall disposition and confidence. On the other hand, high-achieving students living under pressure to keep performing well report less temptation to cheat when they know they’ll have a second chance.

2. By examining test items and students’ performance on retakes, I can often determine whether a student’s low test scores are a knowledge issue or related to the question format. For instance, if a student scores low on multiple-choice responses in all sections but high on other question types, that learner likely needs help in strategizing how to answer multiple-choice items.

3. You may need to convince peers—and students—of the wisdom of retests. Academically elite students sometimes object to a retesting system because they have become
protective of systems that only value those who score well on an initial test. In terms of convincing colleagues, I’ve found that educators who object to retests have considerable difficulty coming up with any examples of assessments in the “real world” that don’t have a retesting component.

Since I reshaped my testing procedures, I’ve looked into the assessment literature and realized that many researchers conclude that the kind of changes I’ve made increase students’ involvement, achievement, and motivation.² I’m glad I’ve seen it with my own eyes.

Endnotes

“How I Broke My Rule and Learned to Give Retests” by Myron Dueck is reprinted from Educational Leadership, Volume 69, Number 3 (November 2011).

Myron Dueck is a vice principal and teacher in School District 67 in Penticton, British Columbia. He presents frequently on grading and assessment procedures; 250-770-7750; myrondueck@gmail.com.

Copyright © 2011 by ASCD
Learn more about how ASCD supports educators who are making a difference and ensuring students’ success.

www.ascd.org

Requesting Permission
1. For photocopy, electronic and online access, and republication requests, go to the Copyright Clearance Center.
2. Enter the periodical title within the “Get Permission” search field. To translate this article, contact permissions@ascd.org