November 2011 | Volume **69** | Number **3**
Effective Grading Practices Pages 72-75

How I Broke My Rule and Learned to Give Retests

Myron Dueck

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?[1] 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:

- Knowledge Targets: What do I need to know?
- Reasoning Targets: What can I do with what I know?
- Skill Targets: What can I demonstrate?
- 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 assessment 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 separating each test into sections by type of question—multiple choice, short answer, long answer—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:

- The United States in the 1920s: 11 points

- Causes of the Depression: 4 points
- FDR's efforts to end the Depression: 5 points
- Reactions to FDR's New Deal: 7 points
- 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

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 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:

- 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.
- 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.
- 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

[1] Stiggins, R. J., Arter, J., Chappuis, S., & Chappuis, J. (2004). *Classroom assessment for student learning: Doing it right, Using it well*. Portland, OR: Assessment Training Institute.

[2] O'Connor, K. (2011). *A repair kit for grading—15 fixes for broken grades* (2nd ed.). Boston: Pearson.

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**.
