

[edutopia.org](https://www.edutopia.org)

The Evidence-Backed Grader

Youki Terada

10–12 minutes

Let us set the scene: A group of teachers sit at a broad conference table, reading student essays together. One scans an essay and gives it a C, noting its lack of coherence. Another pushes the same essay back and pronounces it a B minus, pointing out that the author used quotes well and included insightful analysis.

Sound familiar? For English teacher Seth Czarnecki, this was a quadrennial ritual at his Massachusetts high school. “Though the process varies from time to time, the results are the same. Some of us obsess over pluses and minuses. Others plumb the sample for deficiencies. Ultimately, we leave the exercise in disagreement about the grade the essay deserves,” Czarnecki wrote in [English Journal](#) earlier this year.

If grading were merely imprecise but still highly motivating for students, it might justify placing an even greater emphasis on traditional assessment practices. But that’s not the case, says Chris Hulleman, a professor and researcher at the University of Virginia, and an expert in student motivation. “Despite the conventional wisdom in education, grades don’t motivate students to do their best work, nor do they lead to better learning or performance,” he wrote in [an Edutopia article](#) coauthored by science teacher Ian Kelleher.

This isn’t about throwing grades out entirely. Some method for measuring student knowledge at regular intervals—using standards-based grading, portfolios, or student conferences, for example—is needed to provide stakeholders with a window into academic progress. And there are forms of grading, such as multiple-choice tests and single-answer mathematical exams, that are more precise than the example Czarnecki provides. In the end, solutions to the problem of grading do not need to be absolute. Teachers can still rely on periodic summative assessments and consider evidence-backed ways to reduce the demotivating impact of grading, create more precise measures, and prioritize the messy process of learning over the largely artificial cleanliness of grading.

MIND YOUR ZEROS

A through F letter grades and the 100-point scale feel like eternal verities—systems handed down from the heavens, fully formed. In reality, the 100-point grading scale made its U.S. debut nearly two centuries ago and was originally centered around the 50-point mark, with scores rarely reaching the upper and lower extremes, according to a [2013 study](#). For your grandparents’ grandparents, then, a score of 0 for missed work would be a setback, but not an insurmountable one. Today’s version of the 100-point grading scale, however—after shifting upward to align with the A through F grading scale—is a “badly lopsided scale that is heavily gamed against the student,” the researchers concluded. Factor a single zero into a relatively strong quarter of learning, and a previously A student may never fully recover.

Handing out stiff sentences for missing work, some educators argue, sensibly, teaches students important lessons about accountability and prepares them for real-world consequences. But a [survey from 2022](#) reveals that extensions are frequently granted in professional settings, and in a [2012 study](#), researchers discovered that when the minimum mark in school was a 50 instead of a zero, students put more effort into their learning, earned higher test scores, and graduated at higher rates than their peers under traditional grading schemes. Severe grading practices, the researchers explained, can trigger “defensive and self-destructive responses in students” that can hamper motivation and draw out disruptive behavior.

It might still make sense to give zeros under some circumstances. But the research suggests that it’s better to look for opportunities to give students a path forward. Simple mathematical adjustments, such as dropping the lowest grade (or both the lowest and highest grades) can remove anomalous scores, improve student motivation, and provide a more accurate picture of a student’s ability.

HOLD YOUR CARDS

You can maintain report cards and some, or even most, of your grading practices but find innovative ways to prioritize process over product. In a [2021 study](#), researchers proposed a simple tweak to the grading sequence.

Undergraduate students were randomly assigned to first receive either grades or written feedback on their lab assignments. Those who saw their feedback before the grade became more proficient learners, outperforming their peers by a full two-thirds of a letter grade on future assignments. “Prioritizing written teacher comments can support students to understand their strengths and weaknesses, allowing them to allocate effort to aspects that need improvement. This important process can be undermined by seeing a grade,” the study authors concluded.

To cultivate an atmosphere that encourages creativity and curiosity, teachers at King Middle School in Portland, Maine, make it a point to delay grades until the end of the unit—a mistake-friendly strategy that motivates students to be creative and take intellectual risks. Emphasizing grades too early in the learning process can derail students, explains English language learners teacher Kirsten McWilliams, but delaying grades “is a great way to just give them fluency and comfort with the writing process.”

GO LOW-STAKES, FREQUENTLY

Quizzes, often thought of as a quick way to measure knowledge, are surprisingly flexible tools. A deep body of research reveals that they improve learning, too—an unexpected benefit often referred to as the *testing effect*.

Repeated quizzing tends to work wonders. A [2013 study](#), for example, demonstrated that quizzing students frequently while providing corrective feedback significantly improved learning outcomes—an effect that was still detectable five weeks later. There’s no need to invest a ton of teacher time, either, since even simple quizzing formats seem to do the job. A [study from 2014](#) looked at the impact of short-answer and multiple-choice quizzes on middle school students and concluded that “frequent classroom quizzing with feedback” dramatically outperformed rereading and restudying on learning outcomes—and that “multiple-choice quizzing [was] as effective as short-answer quizzing for this purpose.”

Lowering the stakes also means lowering blood pressure: A [2014 study](#) demonstrated that breaking bigger tests into smaller retrieval sessions reduced final test anxiety for 72 percent of middle and high school students. To change students’ mindsets around testing, consider calling your low-stakes sessions “practice” rather than “quizzes,” and use digital tools like Kahoot or Quizizz to speed up the process, allowing you to see the results in real time and even to gamify your quizzing.

PEER GRADING (WITH TRAINING WHEELS)

Fair, reliable assessment instruments are hard to design and can be difficult to respond to quickly and meaningfully. In many cases, according to former high school mathematics teacher Kareem Farah, now the founder of the Modern Classroom Project, [assessment becomes the cart that drives the horse](#)—teachers know that practice makes perfect but assign less work because they feel incapable of grading the products.

Recent research suggests that there are real alternatives, if you plan accordingly. In a [2022 meta-analysis](#), for example, researchers looked at 175 studies on self-assessment and peer assessment and found that asking students to take active roles in feedback and evaluation “led to significantly better academic performance” across all age groups.

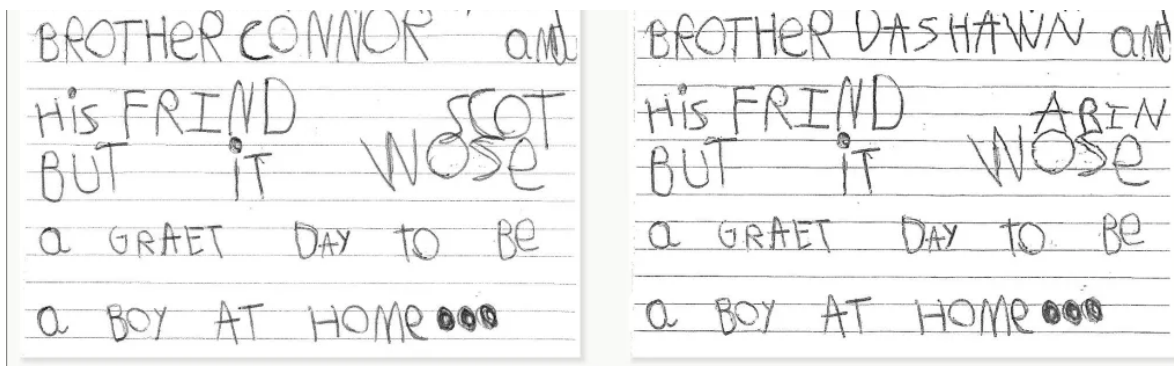
But teachers can’t just ask students to grade and expect big results, the researchers caution. Spend time modeling productive feedback, and provide rubrics, checklists, or exemplars to ensure that students give helpful, learning-oriented feedback. [Research from 2023](#) confirms the finding, revealing that high school students improved their writing by a half-letter grade when they revised while referring to mentor texts or rubrics that laid out expectations, such as narrative cohesiveness or the importance of making a central claim.

LEAN ON RUBRICS

Even with the best intentions, grading can be biased, subtly injecting variability and unpredictability into students’ scores. Research, for example, shows that teachers unwittingly [award higher grades to essays with good handwriting](#), are more lenient toward [boys when they submit partial math solutions](#), and associate [being overweight with laziness and low academic potential](#). Rubrics can help a great deal, research suggests, by providing a structured way to grade subjective work products, reducing the factors contributing to the grade, and explicitly guiding student efforts.

In a [2020 study](#), David Quinn, an assistant professor of education at the USC Rossier School of Education, asked teachers to grade personal essays written by a fictional second-grade student. Two versions of the essay were produced, with one subtle difference: The name of a sibling referenced in the essays was either “Dashawn” or “Connor,” signaling a possible racial difference. Despite being virtually identical, the essays including the name Dashawn were 4.7 percentage points less likely to meet grade-level standards than their Connor counterparts.





Courtesy of David M. Quinn

Nearly identical writing samples—the only difference is the names mentioned—can yield different grades, according to research from USC professor David Quinn.

Bias seeps in where standards are lacking: “If teachers are evaluating student work and they are unsure what standard to compare the work to, implicit stereotypes can ‘fill in the blanks,’” Quinn explained in the study. When teachers used a grading rubric, on the other hand—one that guided teachers to look for specific elements such as being able to recount an event with details—the grading bias was nearly eliminated, he discovered.

To improve your grading, you can use [rubrics that identify clear standards](#) and invite [other teachers to audit](#) your assessment policies and materials, [said](#) Quinn. High school teacher Danah Hashem uses the [single-point rubric](#)—which focuses on a single element, such as “Uses clear examples to support the argument”—to simplify the activity, reduce noise in the feedback, and shift student attention toward a single area of improvement. Teacher Jacqueline Harmer uses rubrics to help her students build [metacognitive strategies](#), reflecting on what they know while planning for their future learning.