

From Data to Decisions: How Data Should Shape Instruction, Not Just Measure It



Page 3

Robert Neubecker for Education Week

EDITOR'S NOTE

Classroom data is most valuable when it **shapes teaching and learning**—not when it sits static on a spreadsheet. Educators are pushing back on standardized test scores and traditional metrics that often fail to capture student growth or guide daily instruction. Instead, the focus is shifting toward **formative, competency-based, and strengths-driven assessments** that offer **real-time insights**. Used well, data supports targeted instruction, differentiation, student support, teacher collaboration, and professional judgment. The goal isn't more data—it's **better data** use to **drive meaningful learning outcomes**.



Vanessa Solis/Education Week via Canva

Page 16

What Teachers Really Think About State Testing.....2

In Case You Missed It: How Schools Are Measuring Student Success.....3

Why Some Schools Are Ditching Class Rank and Weighted GPAs..... 8

What Teachers Should Know About Integrating Formative Assessment With Instruction.....10

OPINION

Teacher Collaboration Often Means Analyzing Student Data to Boost Learning. But Does It Work?..... 11

We Need to Stop Overrelying On Student Test Scores.....14

Principals Often Misuse Student Achievement Data. Here's How To Get It Right.....16

A Strengths-Based Guide To Assessing Student Progress (DOWNLOADABLE).....18

Published May 27, 2025

What Teachers Really Think About State Testing

By Jennifer Vilcarino

As students wrap up the school year, their last hurdle is often state testing, which can spark discussions about whether or not these exams accurately measure student learning.

In the spring, states must test students in grades 3 through 8 and once in high school on reading and math, based on accountability requirements established by the No Child Left Behind Act and reauthorized in the Every Student Succeeds Act in 2015.

However, educators have varied opinions about the effectiveness of these exams. Some believe the end-of-year state assessments measure only a student's ability to take a test and use a one-size-fits-all approach in education.

Oklahoma is testing how flexible ESSA and these requirements are under President Donald Trump's Department of Education. Ryan Walters, Oklahoma's superintendent of public instruction, has said the state is exploring how testing can "look different."

In an informal Education Week LinkedIn poll, a majority of respondents, 84%, voted that state testing is not an appropriate benchmark for student learning, while 11% said it "somewhat" is. Just 5% believe state testing accurately measures student learning.

Educators shared more details about their opinions on state testing in the comment section. Here are some of their responses about state testing, edited lightly for clarity:

Some teachers want less of a focus on tests

Some educators who responded to the poll said state testing only measures a student's ability to memorize and work well under pressure.

"Standardized tests only consistently prove one thing... children in poverty don't test as well. Take the testing money and invest it into the classroom. Smaller class sizes, personalized instruction, and social services support."

—Rachel L.

"It's a test of endurance more than anything. This is just one example of what is wrong with standardized high-stakes testing."

—Marsha L.



Eric Thayer/AP

A teacher points to a board as students listen in a fourth grade classroom at William Jefferson Clinton Elementary in Compton, Calif. State testing happens every spring and educators share their thoughts on whether these assessments accurately reflect student learning.

Testing can be used to evaluate schools, rather than individuals

Other educators said these assessments ensure specific groups, such as English learners, students with disabilities, and students from low-income households, do not fall behind.

"It does not do a good job of evaluating individuals. However, longitudinally over time, it can show trends in different educational systems and their effectiveness in different areas."

—Adam S.

"It is a snapshot in time. It should be used as a benchmark to adjust curriculum, teaching, and learning."

—JoAnn C.

State testing is only one way of measuring student learning

Other educators believe state testing can be one data point among many to illustrate how a student is progressing.

"State tests are designed to bring the bottom up. They make sure students can write an essay and do algebra and geometry. I said somewhat because some students do not do well on exams; thus, they

should be able to show their learning in different ways. State tests are just one of many ways to measure student learning."

—Mark D.

"Has value as one of multiple measures. Assuming the assessments are standards-aligned, they can provide a snapshot at a high level—statewide, regional, institutional. The closer you get to a single student the value declines. Learning is complex, variable, and contextual—your other measures need to account for that."

—Jim B. ■

Published April 23, 2025

In Case You Missed It: How Schools Are Measuring Student Success

By Daniela Franco Brown

The A-F grading scale has long been the cornerstone of measuring success in K-12 schools, but educators and researchers continue to question whether those letters truly reflect what students know, and what they can do.

A recent Education Week special report, “Grading and Assessment: How to Best Measure Student Success,” examines how districts are rethinking everything from extra credit to class rank, and how some teachers are navigating sensitive conversations around grade changes.

As districts nationwide seek to better serve all students and examine the limitations of traditional grading systems, this report explores the challenges and innovations surrounding the quest for a more accurate and equitable assessment of student learning.

What are grades really for? What research says about 4 common answers

When Laura Link works with districts to review their grading policies, she starts by asking a simple question: What are grades for?

Are grades meant to motivate students? To help teachers determine who needs more support? To inform parents of their children’s progress?

Link, an associate professor of education at the University of North Dakota, studies how K-12 schools formulate grades and how those decisions affect student learning. She also consults with school systems about their practices, starting with surveys and focus groups with families, caregivers, teachers, students, and administrators.

Those findings often expose misalignments in expectations about what grades and homework are supposed to accomplish.

Those differing priorities—and the tensions between them—cut to the heart of the grading debate. Grades can’t be all things to all people, assessment experts say. Districts must weigh tradeoffs and communicate clearly when revising grading policies.

Grades “can affect students’ trajectories—



Robert Neubecker for Education Week

not just academic trajectories but life trajectories,” Link said. “We use that symbol to make day-to-day decisions, as teachers and as families, that affect students’ futures.”

Inside one district’s big policy shift on grading for equity

Educators at the San Leandro Unified school district in Northern California had a grading and assessment dilemma.

In 2013, some students getting D and F letter grades were meeting content standards on state and local assessments. Some students getting A’s were struggling on state tests and Advanced Placement coursework. Some students, who educators thought would need an intervention course, didn’t actually need it based on various measures other than grades.

Even after the district shifted in 2014 to allow for project-based learning—for which students demonstrate their content knowledge in a variety of ways—the mismatch between letter grades and test scores persisted, said Sonal Patel, an assistant superintendent of the district. Students taking the same course but switching teachers saw their grades vary depending on which teacher they had.

This prompted district leaders to invest in teacher training and reimagine grading so that each letter grade carries equal weight, an F is not overweighted, and grades are based solely on students’ knowledge of required academic content—excluding factors like extra credit or classroom behavior.

Why some schools are ditching class rank and weighted GPAs

Superintendent Alison Villanueva was about a month into leading Connecticut’s Watertown school district when she suddenly found herself facing a “catastrophe, PR nightmare.”

A high school principal published the names of that year’s valedictorian and salutatorian in a press release and on social media. Then, the district got a call from a parent arguing that the information was wrong and that their own child had, in fact, come in second in the graduating class.

The student whose name the district had publicized actually ranked third.

Though this may seem like a minor mishap to someone who doesn’t work in K-12 education, Villanueva “was in a frenzy,” she recalled four years later. “If you have something like this that’s so monumentally incorrect, it puts doubt in people’s minds about everything you’ve done.”

“**[Grades] can affect students’ trajectories—not just academic trajectories but life trajectories.**”

LAURA LINK

Associate professor of education
University of North Dakota

What should teachers do when students—or parents—ask for a better grade?

What's in a grade?

On its surface, a letter or number grade would appear to be a fixed measure of a student's performance on a given assessment: a test, essay, project, or other tangible piece of work.

Ideally, a grade should “accurately reflect a student's current understanding of the course content, free from biases,” according to Joe Feldman, an educational consultant, former teacher, principal, and author of the book *Grading for Equity*. But what about when students or parents ask, and are granted requests to change grades?

This practice, unflatteringly referred to as “grade grubbing,” has become increasingly common.

How teachers approach grading, in charts

Grades are a well-established part of the K-12 education system that, in most school districts, play a decisive role in students' academic standing.

But what goes into the single letter or number that a grade represents? Just how malleable is a grade after it's been issued? And what effect does grading work have on students' performance?

In December 2024, the EdWeek Research Center conducted a national survey of 759 K-12 teachers to glean answers to these and related questions.

A strengths-based guide to assessing student progress

Teachers use rubrics to ensure consistency and clarity in grading student assignments by carefully detailing what a successful project looks like.

But the seeming simplicity of these assessment tools masks complicated questions about how to measure student progress and how to encourage continued learning, said Kevin Perks, the senior director of Quality Schools and Districts at WestEd, a nonpartisan education research organization.

“You have to determine what constructs to measure; how to measure those accurately, with validity; and then how to develop instruments that can be used consistently,” he said. “Assessment literacy is one of the largest knowledge gaps teachers have. It's an area where they aren't really trained.”

When designed effectively, rubrics can clarify expectations, minimize subjectivity, and standardize grading criteria across multiple teachers, said Perks, who leads training sessions for teachers about classroom assessment.

There are many schools of thought about how to properly design these grading tools. ■

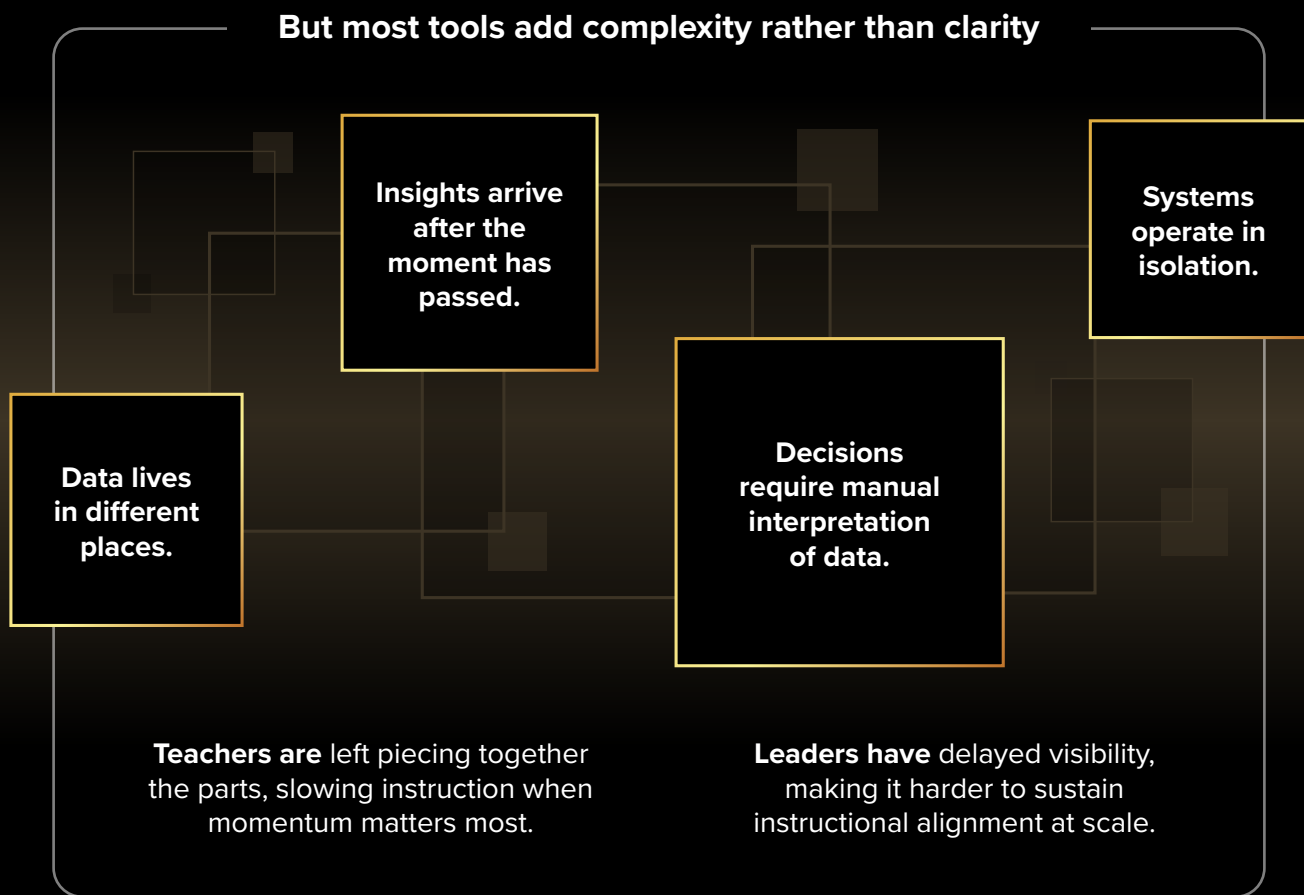
Additional Resource
View this article's charts



The challenge:

Teaching often outpaces technology

An effective multi-tiered system of support (MTSS) requires fast, confident instructional decisions grounded in real performance data—not static pathways or disconnected tools. Educators need insights into what students need *now*, along with the ability to act in the moment while staying aligned to standards and curriculum.



Schools need technology that keeps up—not more layers.

Renaissance

The solution:

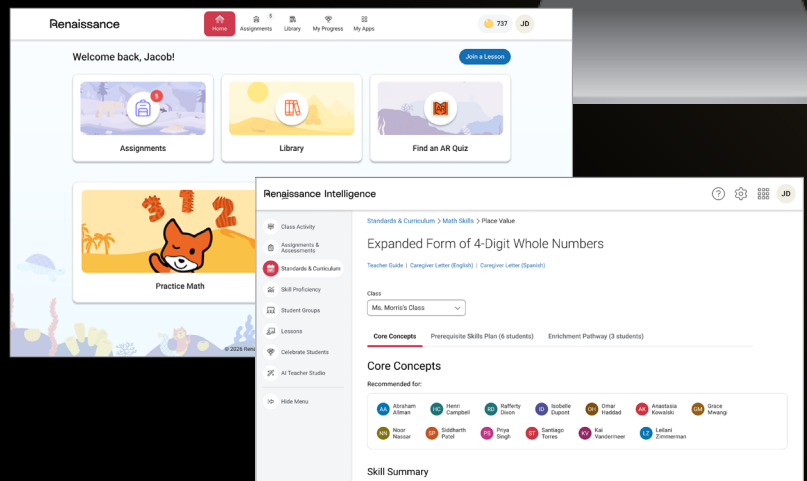
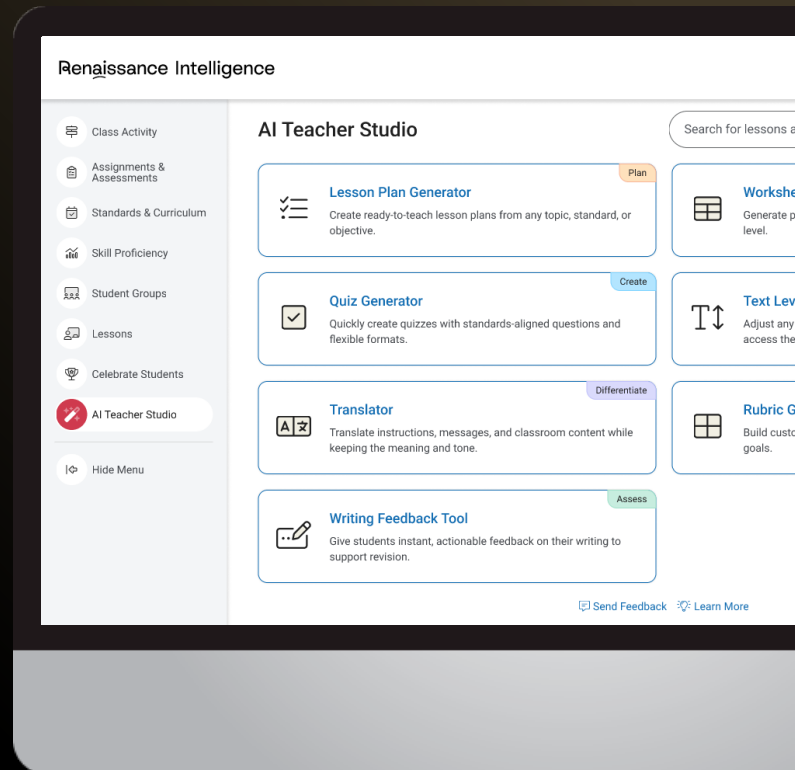
Renaissance IntelligenceSM

Renaissance Intelligence is the groundbreaking new system that unifies assessment, instruction, practice, and curriculum alignment. It adapts to classroom context, promotes instructional coherence, and connects directly to the skills students need next.

Powered by decades of learning science and research-backed data, Renaissance Intelligence delivers:

- Unified assessment-instruction-practice workflows
- Adaptive AI that uses classroom context to surface insights and recommendations
- Deep alignment to standards, skills, and core curricula

The result: Every decision stays connected to what students need next.



Renaissance Intelligence connects the loop from assessment to guided instruction and targeted practice—so insight becomes action when it matters.

See the impact Renaissance Intelligence can have in your schools

Learn more

Renaissance



Robert Neubecker for Education Week

Published April 14, 2025

Why Some Schools Are Ditching Class Rank and Weighted GPAs

By Alyson Klein

Superintendent Alison Villanueva was about a month into leading Connecticut's Water-town school district when she suddenly found herself facing a "catastrophe, PR nightmare."

A high school principal published the names of that year's valedictorian and salutatorian in a press release and on social media. Then, the district got a call from a parent arguing that the information was wrong and that their own child had, in fact, come in second in the graduating class.

The student whose name the district had publicized actually ranked third.

Though this may seem like a minor mishap to someone who doesn't work in K-12 education, Villanueva "was in a frenzy," she recalled four years later. "If you have something like this that's so monumentally incorrect, it puts doubt in people's minds about everything you've done."

The mistake resulted in more than a corrected press release. After some fact-finding and soul searching, the district made significant changes to the way it calculates weighted GPAs and how it decides which students can take the challenging courses that help boost those averages.

The whole scenario also got Villanueva thinking about the logic behind weighted GPAs and class rank.

"We're trying to calculate our way to the top of high school," she said. "What's the point of that? Just [to] have the social aspect of being deemed the valedictorian, a salutatorian?"

To be clear, Villanueva's skepticism isn't meant to diminish high achievers, she said. But she and other educators wonder if a system that emphasizes—and publicly rewards—grades over mastery of skills and knowledge will really prepare students for the future workforce.

After all, artificial intelligence tools can pass just about every Advanced Placement test—the capstone assessment for the courses that often carry the biggest bang for the weighted GPA buck.

Such questions are especially fraught given that low-income students and students of color are less likely to be steered toward rigorous courses and therefore earn top honors in schools that weight GPAs, experts say.

What's more, even as K-12 writ large pushes past the "college for all" ethos of past decades, most schools' methods for weighting GPAs and calculating class rank offer little benefit for students in career-and-technical education.

"There's one definition of academic success, and it's through these weighted courses," said Kristin Klopfenstein, the director of the Colora-

do Evaluation and Action Lab at the University of Denver who has conducted research on the topic. "That reflects your local values, whether you recognize it or not."

Schools' methods for weighting GPA differ significantly

Though most high schools weigh GPAs, there's wide variation in how those averages are calculated, including exactly which courses carry extra points, Klopfenstein said.

Some schools, for instance, give extra credit for only college-level courses—such as AP, International Baccalaureate, or dual enrollment, she said. Others offer a GPA bonus for those classes plus all courses labeled "honors." Among those that give credit for both college-level and honors, some give students more of an advantage for taking college-level classes. Others handle AP/IB classes and honors the same.

There's been little examination of which approach is best, Klopfenstein said.

"The only thing that's consistent is that there's no consistency in how schools treat grade weighting," she said.

Weighting GPAs does indeed have a modest, positive influence on the number of advanced courses students take, Klopfenstein found. But the effect is limited primarily to white students who are not eligible for free or reduced-price lunch.

That may be because many schools aren't directing students from certain racial backgrounds to the rigorous courses that may lift their weighted GPAs, according to separate research conducted by Sarah Morris, the research liaison for Rice University's Houston Education Research Consortium.

Morris found that Black students in Arkansas are three times less likely to be placed in advanced courses as Asian students, even when they have the same academic ability.

"That makes me feel really terrible about weighted GPAs, because we're talking about a whole group of students that just kaboom [can't take advantage of] those opportunities," she said.

Educators are divided on whether grades themselves motivate students—much less the race for the highest weighted GPA or the valedictorian slot.

About 1 in 6 educators—16 percent—say that students would be more academically motivated if there was less emphasis on grades, according to an EdWeek Research Center survey of 759 teachers conducted in December.

Nearly a quarter—23 percent—believed student outcomes would improve if grades were

emphasized more. Another quarter says that the current level of emphasis maximizes student engagement. And a plurality—36 percent—believe that nothing about student motivation would change if grades were emphasized differently one way or the other.

It's typically only a small handful of high achievers scrambling to have the highest weighted GPA or become valedictorian, said David Lang, the chair of the economics department at California State University, Sacramento, who has studied GPAs.

"The rising tide, in this case, does not raise all ships. It just raises the ships that are at the top," Lang said. "The other ones just kind of look at this as a game that they don't get to participate in. No kid that is 312 out of [a class of] 600 is trying to battle for 311."

The class-rank race carries unintended consequences

Some schools are responding to the fact that only a few students are motivated by class rank by reworking how they order students or weight their GPAs.

Virginia's more than 4,000 student Winchester school district has switched from traditional class rank to designating students as summa cum laude, magna cum laude, and cum laude, just as colleges do—a change the school board approved on the recommendation of Superintendent Jason Van Heukelum.

Over his decades in education, Van Heukelum has watched students—including his own daughter—make "dumb decisions" to protect their GPAs, he said.

Some top-notch students are much less likely to take electives that don't carry any extra points. They've figured out that even acing an art course, for instance, could drag down a GPA above a 4.0, since art classes don't carry the extra points that some academic courses, such as AP, do.

"If you're really obsessed with your GPA, then you're going to nickel and dime that and you're going to make bad choices as a 16-, 17-year-old," Van Heukelum said. "You're not going to really enjoy the offerings that your high school has available."

For instance, an aspiring engineer who is striving to stay at the top of their high school class may opt to take AP Psychology, even if they "have no interest in the field," because it gives them a GPA quality point. And they may eschew a welding course, which has plenty of application to engineering, because CTE classes don't offer a weighted GPA boost.

In fact, ditching the practice of weighted

GPAs altogether is worth considering, Van Heukelum said. Given the wide variance in how schools approach the averages, he's not sure how meaningful it is to college-admissions officers or potential employers.

But that shift may make parents anxious, he acknowledged.

"I think that's the fear factor for families in particular is that our kids aren't as competitive [as they could be if] we don't weigh enough," Van Heukelum said.

California Area school district, a 900-student rural district in Pennsylvania's Allegheny mountains, is adopting a similar approach—designating top students as summa, magna, and cum laude—and will eventually use that system in lieu of class rank, said Superintendent Laura Jacob.

The shift will be implemented gradually, so that students don't feel caught off guard, she said.

The district already allows students to opt out of grades entirely in favor of a competency-based approach that results in a mastery transcript.

For kids who stick with letter grades, the traditional ranking system "is damaging to our top 10 kids that are in that class-rank battle," Jacob said. "Kids are literally breaking down, balling because they got one B on a test."

The district is trying to instill values like grit and a growth mindset in its students, she said. If students are rattled by a single less-than-perfect score, "we're not preparing [them] for the challenges of the world when [they] leave high school," Jacob said.

Experts advise schools to revisit their weighted GPA policies periodically

In Watertown, the wrong salutorian blunder prompted Michael Dalton, the district's instructional technology specialist, to look under the hood at how the district calculates its weighted grade point averages.

He found that the district's system rewarded students who took challenging math and science classes but didn't offer as big a benefit to those who enrolled in high-level humanities courses.

At the same time, Villanueva and other leaders realized fewer students of color were taking AP classes and other rigorous courses that could boost their GPAs.

"So, we're thinking, 'What the heck is happening there?'" Villanueva said. "Because that's not fair, either."

The district made a policy change, putting

challenging humanities classes on equal footing with tough STEM courses in calculating weighted GPAs. And it made it easier for students from a variety of academic backgrounds to enroll in advanced courses.

In the past, "there had been such gatekeeping that you only [had] the [highest achieving] kids in your AP classes, but now we are essentially increasing the number of students of all races and ethnicities and backgrounds and abilities to go into these classes," Villanueva said.

It's smart practice for districts to periodically take a fresh look at how weighted GPA and class rank are calculated, Klopfenstein said.

"I think sometimes it's legacy. The decisions are made, then it just becomes 'this is how we've always done it,'" she said.

Instead of relying on old systems, districts should consider revisiting their policies every few years, she said. They should ask questions like: "Why do we do this this way? Have our values changed? Are there unintended consequences that we have not paid enough attention to?" Klopfenstein said. "We need to be continuously more thoughtful about this." ■



E+/Getty

Published April 17, 2024

What Teachers Should Know About Integrating Formative Assessment With Instruction

By Sarah D. Sparks

Philadelphia—

Teachers need more support to move testing from a “necessary evil” to a classroom tool, experts say.

While summative assessments—like unit quizzes or annual state tests—are used for evaluation and accountability, research shows formative assessments—like puzzles, projects, and class error analyses—can help teachers and students identify misunderstandings and reflect on students’ progress as they’re learning.

While most teachers use at least some formative measures in daily classroom practice, experts at the American Educational Research Association conference here last week argued they need more support to integrate daily assessment practices with overall classroom instruction.

“We need to think about assessment more holistically,” said E. Caroline Wylie, a senior associate at the National Center for the Improvement of Educational Assessment and co-author of the National Academy of Education’s new report on assessment released last week. “Certainly we’ve got to be sharing learning goals for students, assessment, and learning in ways that are recognizable.”

For example, in one study, Dustin Van Orman, a STEM education research associate at Western Washington University, asked a national sample of more than 100 elementary student-teachers who participated in simulations of English, science, and math classes. The teachers were asked how much experience they had in using formative assessment, and then were asked to use information about students in the simulated classes to plan tasks and other formative assessments over three class periods.

About 1 in 5 of the preservice teachers had little to no prior training in formative assessment, and a third had experience only with formal testing, rather than informal measures through tasks and activities. One preservice teacher reported that her mentor-teacher downplayed the usefulness of classroom assessment and “only likes to do assessment if it’s like something she kind of has to do.”

While a majority of preservice teachers in the study could identify tasks students should be able to do as they learn particular academic content, Van Orman and his colleagues found many student-teachers did not set learning goals with their students or set criteria for tests and tasks based on learning goals. Rather, tests and tasks often could be disconnected from overarching instruc-

tional goals. Teachers less experienced in assessment also tended to give more static feedback—praising or correcting students—but not providing information on which students were expected to act.

How to use formative assessments effectively

“Assessment shouldn’t be dropped in from the sky, disconnected from student experience,” said Wylie, who was not part of the preservice teacher study. It’s also important for teachers to understand their own and students’ cultural backgrounds when designing assessments, she noted.

Van Orman and Erin Riley-Lepo, a visiting assistant professor at the College of New Jersey, have been working with researchers, teachers, and principals to develop a framework for teachers’ own assessment literacy.

To effectively use formative assessment in class, they recommended:

- Teachers and students develop a shared understanding of learning goals.
- Assessments focus on what students are learning in ways that the students can recognize, so that they can understand their own progress.
- Students should engage in self- and peer-assessments, to help take ownership of their learning.
- Assessments should make students’ thinking visible to both the teacher and students, to correct misconceptions and build on students’ strengths.
- Assessments should directly inform teachers’ instruction.

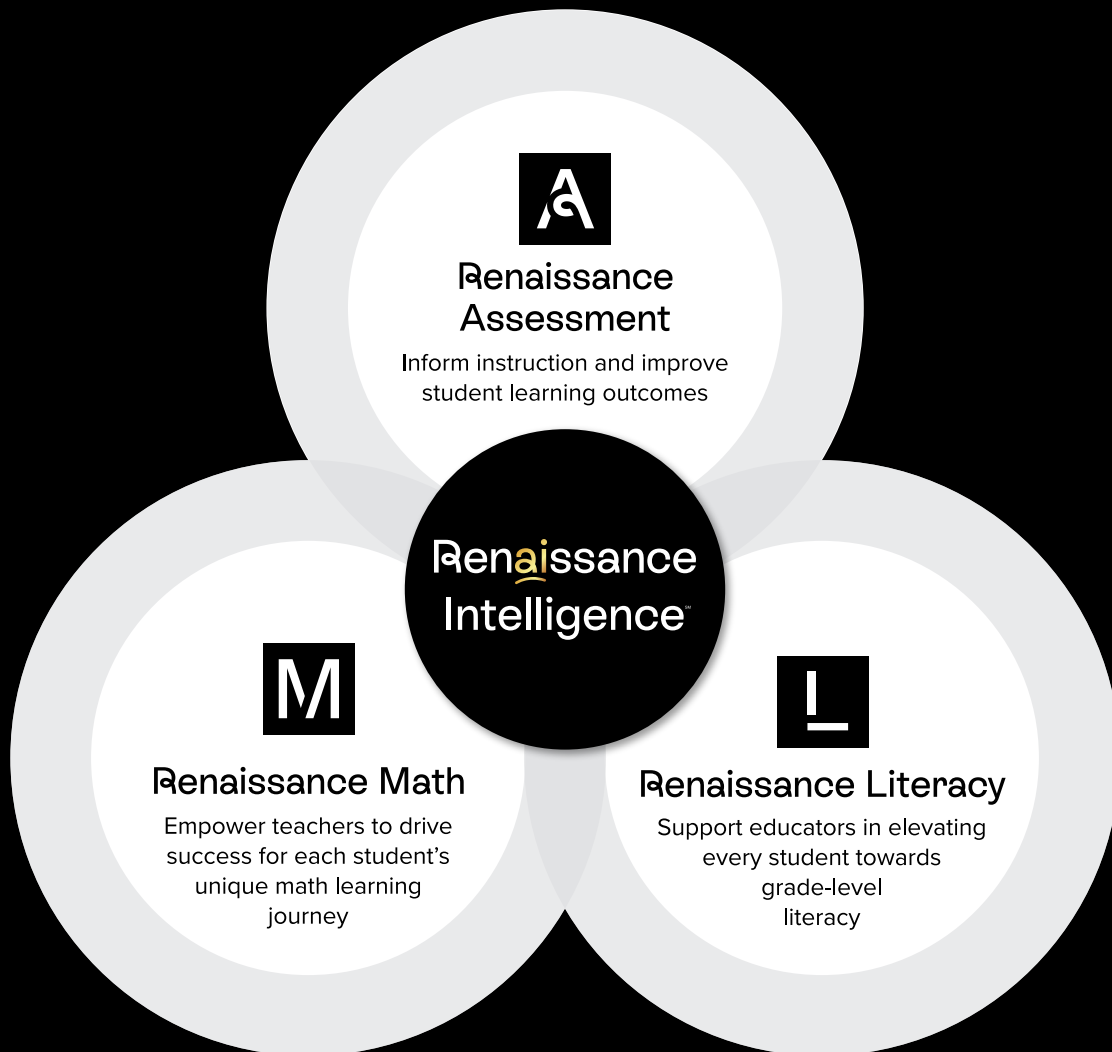
Building better classroom assessment practice also requires support from principals and district leaders.

“We talk about teachers, but the reality is, teachers are embedded in schools and districts with particular approaches and constraints on their assessment that may or may not be—and often is not—supportive of active and focused assessments,” Wylie said.

For example, Wylie noted that in the last four months, she and her research team repeatedly had to cancel professional development for teachers because the principals could not secure enough substitute teachers to cover their classes. ■

One system. Three connected solutions.

The new Renaissance Intelligence is built on tools that educators know and trust—now connected into one system to support data-driven instructional decisions.



Reach out to see how these solutions will
supercharge MTSS in your schools

[Connect with an expert](#)

Renaissance



Z.wei/iStock/Getty and Education Week

OPINION

Published March 27, 2024

Teacher Collaboration Often Means Analyzing Student Data to Boost Learning. But Does It Work?

Teachers deserve a blame-free, action-oriented protocol

By Ronald S. Thomas

When teachers have used scarce professional learning time to analyze student-assessment data, achievement results have often been mixed. And that's despite the fact that teacher collaboration itself has a record of improving student learning. As Harvard scholar Heather Hill noted somewhat delicately in 2020, "The research in this area suggests that district and school leaders should rethink their use of state and interim assessments as the focus of teacher collaboration."

Should school leaders throw out the practice? I can understand why that would be tempting not only from the research findings but from personal experience.

As a data coach for over 30 years, I still too often see teacher teams spend lots of time giving reasons for poor performance that are way outside their ability to have much impact, such as a difficult home life, students' failure to study, or the unfair demands society puts

on schools. There is little attention to students' assets and their specific misunderstandings or to improving future instruction. No wonder analyzing student-assessment data hasn't lived up to what some would call its hype but I would call its potential.

Data analyses often fail to have positive impact on student learning because district directors of assessment (I was one) and university faculty members (I am one) have not given educators the words or structure needed to redirect conversations to what the data show that students know, what students do not know, and what the adults on the team are going to do about it. In other words, leaders have too often told teachers to conduct data analyses but very seldom taught them exactly how to do it.

In her seminal 1982 study "Workplace Conditions of School Success," sociologist Judith Warren Little found that in successful schools, teachers engage in "frequent, continuous, and increasingly concrete and precise talk" about "what teachers do, with what aims, in what situations, with what materials, and with what apparent results." But "concrete and precise talk" does not occur automatically. Here are a few

lessons the teachers in my educational leadership class and I have learned to move teams to a comfort level of being concrete and precise with one another about instruction.

First, a protocol for the conversation must be in place. Using "data dialogue" protocols helps school teams navigate difficult and uncomfortable conversations by providing a predictable series of questions or topics for members to discuss sequentially. Predictability builds teachers' trust with the school leader who could be part of the conversation and with each other.

There are several good protocols in use. Some of the best are based on the premise that teachers are not responsible for their data but are responsible for their response to their data. My leadership students report that this slight wording change makes a world of difference in teachers' attitudes and their commitment to their work.

Effective protocols are action-oriented, with predetermined facilitative moves that avoid the "blame game," triangulate test scores with teachers' and students' observations (sometimes called "street data" or "local expertise"), and require teachers to follow up in their classrooms and report results at the next meeting. Classroom follow-up is often the weakest part of the process, my students say.

In the most helpful protocols, teams address classroom patterns of student performance first. The idea is to avoid fixating on the needs of individual (such as "watch list") students too soon. Once teams head down the path of considering individuals, there is usually no end to the list, collaborative time runs out, and only a few students get attention, while team members whose students aren't being discussed zone out.

My leadership students report that the best protocols aim for team consensus on the most effective, evidence-based responses—the type, timing, and extent of the interventions. Is the concept that was missed by many students so important that new instruction must stop right away and the concept retaught? Or should teachers integrate additional instruction on the weak content seamlessly over time into the next topic? Do students have a basic understanding of the concept but need more practice? Or is the next step reaching out to students ready for enrichment and to those who need additional time on the weak skills?

To empower dialogue participants to act, leaders must grant teams the autonomy not to be bound by standardized pacing charts that do not provide teachers flexibility for differentiated instruction. If a concept is essential for further learning and teachers have multiple data points indicating a widespread need and have identi-

fied a different way to teach it, it is, in my view, educational malpractice not to address it.

Once there is team consensus on the most appropriate follow-up, team members use their combined expertise to help each other plan instructional strategies, in as much detail as possible, so individual teachers don't have to do this on their own time.

Lack of closure is often the missing ingredient to make meetings impactful, according to management expert Paul Axtell. Therefore, effective protocols include a three- to five-minute closing segment to validate individuals' contributions, review the takeaways, and agree on what each person will do and by when.

As teams use practices such as these, they move toward concrete and precise dialogue. Members feel empowered, gain instructional insights from each other, and increase their respect for the protocol and for their professional colleagues. And when teachers return to the next dialogue session with evidence of student improvement, their positive feelings about collaboration and the value of data analysis grow even more.

So throw out team analysis of student-assessment data? Not so fast. School leaders have found that the practice works for teachers and students when teachers are provided with the structure and the autonomy to do it well. ■

OPINION

Published April 28, 2026

We Need to Stop Overrelying on Student Test Scores

By Larry Ferlazzo

What educator hasn't sometimes felt frustrated at the overreliance on test scores to measure the success of their students and their teaching?

'Expanding Rigor'

James Soland is an associate professor of research, statistics, and evaluation at the University of Virginia School of Education and Human Development whose work focuses on assessment (psychometrics), evaluation, and data use:

Educators know the drill: A new program rolls out, someone gathers test scores before and after, and they are told whether it “worked.” But what does that really tell us? Does it help improve teaching? Does it help educators and policymakers understand why it worked here but not there? Does it help us understand what students actually gained and experienced?

As I discuss in a recent blog at the Brookings Institution, right now, the field of education evaluation is fixated on one question—“what works?”—narrowly defined as whether a program causes measurable improvements in things like test scores. That is certainly valuable at times, but it misses too much of what matters in schools. It privileges what we can easily measure over what we ought to understand. And it treats school contexts as interchangeable backdrops rather than vital elements of success.

As teachers and school leaders, you know that learning is messy, local, and human. It unfolds in specific classrooms with specific students and adults doing real work together. The current dominant approach—evaluating programs like black boxes and judging them by a narrow set of outcomes—doesn't capture that reality. It leaves out teacher expertise, student experience, school culture, and the many contextual conditions that make a strategy effective here but not somewhere else.

We need evaluation that learns from you—the educators in the trenches—not just about students' test results. You are the people who see

when a strategy sparks student curiosity, when it bumps up against local realities, or when it might not work when applied to a different set of students. You also know when a program is truly moving the needle, versus being done purely out of compliance.

What's the problem with “black box” evaluations?

Most rigorous program evaluations focus on isolating causal effects: “Did X cause Y?” That's mostly done with experiments or quasi-experiments using standardized outcomes that policymakers and researchers can compare across time and place. But this approach has two big limitations:

It favors outcomes that are easy to quantify—like state test scores—over equally important outcomes that are harder to measure, like critical thinking, collaboration, or students' sense of belonging. Those latter components are often at the heart of teachers' daily decisions and matter deeply for long-term learning.

It treats context—the school, community norms, teacher skills, resources, and culture—as something to control away, rather than a source of insight about how and why something works.

This narrow focus leads to evaluations that feel detached from reality. They tell you whether something worked somewhere but not why it worked, how it worked, and under what conditions it might work in your own context.

So what might a better evaluation look like? Let's use an example of a socio-emotional intervention (such as to boost growth mindset or self-management skills) designed to improve that socio-emotional competency and, thereby, also improve achievement. Here are some key elements:

1. Broaden outcomes beyond test scores.

Standardized tests capture important academic skills, but they miss socio-emotional growth, critical reasoning, cultural competence, and other dimensions of learning that teachers nurture every day. When evaluation counts these too—even if they're harder to quantify—it aligns more closely with what matters for students.

In the socio-emotional-competency ex-

ample, it would mean not only looking at achievement gains but also at changes in self-management or growth mindset, ideally using a survey measure designed to understand change over time. It would further involve asking teachers whether they think the intervention actually improved the competency or if it was more likely a measurement artifact (e.g., students better anticipating the “correct” answer on the survey after the intervention).

2. Mix numbers with narratives.

Rigorous causal work has its place—but it should sit next to rich qualitative evidence. That means intentionally gathering teacher perspectives, student voices, and descriptions of administrator experiences. Qualitative research has often been peripheral in program evaluation, but it helps us understand mechanisms—the how and why of what works—not just the if. In the case of the socio-emotional-competency intervention, interviews with teachers would ask if the teacher felt there was a valid, causal chain where the intervention increased self-management or growth mindset and that improvement then caused achievement gains.

They would talk about whether the intervention is easy enough to implement that it could be part of common practice. If the intervention did not show gains (in the competency or achievement), teachers would provide qualitative data on why not.

3. Make context part of the question, not something to control away.

Instead of treating local conditions as noise, good evaluation treats them as data. Knowing how a rural school engaged parents or how a multilingual classroom adapted a reading program can teach us about transportability and adaptation.

For socio-emotional conferences, that could look like asking teachers whether they felt there was support for the intervention (e.g., sufficient time to implement it well), if there were bureaucratic hurdles, if it did or did not work for students with particular learning challenges (e.g., students with a particular IEP), how their particular school and setting affected outcomes, etc.

4. Use mechanisms to guide improvement.

Instead of only reporting that an intervention “worked,” evaluations should articulate how it produced results. Was it because teachers had more collaboration time? Because students engaged more deeply with texts that reflected their lived experience? Because instructional coaching supported risk-taking? Because teachers recognized the value and bought into the strategy? These mechanisms—not just outcomes—are the critical lessons for replication and improvement.

In the socio-emotional-competency case, all of these mechanisms could emerge during teacher interviews, surveys, focus groups, or whatever was the most efficient use of their time. (Obviously, these additional data would be collected mainly in large-scale evaluations with sufficient resources to compensate teachers fairly.)

Putting contexts and conditions front and center with help from teachers

Teachers are constantly evaluating: They watch a lesson unfold, notice signs of engagement, troubleshoot misconceptions, and decide what to try next. That expertise—grounded in context and informed by deep knowledge of students—needs to be part of how we study educational effectiveness. By expanding our definition of evidence and privileging why and how as much as whether, we make evaluation more useful.

And I believe we can, by expanding the aperture of the questions we ask, just maybe, move past a hyper-fixation on test scores.

Moving beyond test scores and black boxes doesn't mean abandoning rigor. It means expanding rigor to include modes of inquiry that respect complexity without losing clarity. It means building evaluation systems that help educators and policymakers learn what works for whom, why it works, and what to do next.

Thanks to James for contributing his thoughts. ■

Larry Ferlazzo is a former award-winning high school English and social studies teacher of more than two decades. He is currently a volunteer tutor to English-learner newcomers at a local school and to youth in juvenile hall.

Renaissance Intelligence



- Class Activity
- Assignments & Assessments
- Standards & Curriculum
- Skill Proficiency
- Student Groups
- Lessons
- Celebrate Students
- AI Teacher Studio



Give them a helping
hand, **before they**
raise theirs.

The full picture changes everything.

Built for how classrooms actually work—and designed to keep teachers at the center of instruction.

See what's possible

Renaissance



Vanessa Solis/Education Week via Canva

OPINION

Published October 28, 2025

Principals Often Misuse Student Achievement Data. Here's How To Get It Right

8 recommendations for dodging common data sins

By David E. DeMatthews & Lebon "Trey" D. James III

Each school year, one important task on a principal's plate is to review student-achievement data to identify trends and adjust goals. However, several common misinterpretations and misuse of such data from standardized assessments can have serious consequences for teachers, staff, and families. In our own experiences in administration and school leadership research, we have seen firsthand how novice and veteran principals fall short of leveraging the valuable data from state tests.

The reality is that many principals have been socialized to use data in problematic ways. Others might work for supervisors who are misinformed on effective data use.

Some common ways we have witnessed principals misuse data include permanently labeling particular students as "low performing" and tracking them into less rigorous or remedial courses focused solely on test preparation, penalizing teachers or students for test scores

without any consideration of context or growth over time, pressuring teachers to teach to the test, sharing identifiable student data on publicly displayed data walls, and ignoring factors like students' language, disabilities, or home-life circumstances when prioritizing resources to raise student achievement.

These approaches to data use are harmful to school culture and teacher working conditions, and they do not support long-term sustainable improvements to student achievement. They can lead to misdiagnosing students in need of interventions or special education supports, which further undermines student morale, causing increased stress, disengagement, and dropout.

Testing companies, research centers, and researchers have long clarified that standardized tests should not be used as a sole or primary measure of student achievement, teacher effectiveness, or school quality. These scores reflect only a snapshot, not deeper learning or critical-thinking skills a student or student body might possess through engaging with high-quality curriculum and instruction.

Perhaps, most important, standardized-test

scores must be understood not only as important measures of student progress but also as imperfect. What is tested and how test results are interpreted are political decisions that have implications for funding decisions, school closures, teacher evaluations, and the broader stigmatization of schools serving historically marginalized communities.

Despite these challenges, principals can effectively utilize standardized-testing data to make meaningful improvements on their campus, especially if they are educated on the appropriate uses of data.

We offer a few recommendations and reminders to principals as they dig into their achievement data with their administrative and instructional partners:

- Use standardized-test scores alongside other forms of data collected during the year, including observations of teacher lessons and student work, classroom interventions, and assessment data, as well as attendance and behavioral data. We recommend reviewing the AASA's "Using Data to Improve Schools: What's Working" and the National Forum on Education Statistics' "Data Use for School and District Leaders" reports.
- Analyze data based on campus goals and student populations with teams that include diverse perspectives.
- Pinpoint strengths and areas of growth. For example, consider where students performed well and where they struggled.
- Identify students who might benefit from specific types of interventions, supports, and enrichment opportunities, including those with disabilities and English learners.
- Disaggregate data to identify and address potential disparities, particularly for students of color, students with disabilities, English learners, and students who are struggling academically.
- Plan targeted areas of professional development for all teachers, teacher teams, and individual teachers as appropriate. Both the Institute of Education Sciences and the National Association of Elementary

School Principals have produced free, time-tested guides on using student-achievement data to support instructional decisionmaking.

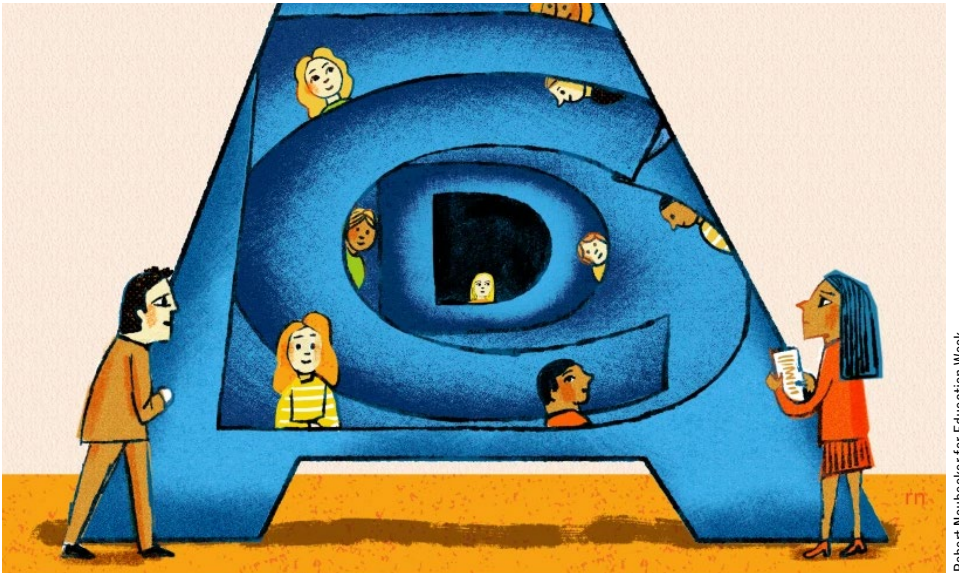
- Determine where additional support staff is needed and adjust the master schedule to enable students to receive appropriate instruction and interventions aligned to multitiered systems of support. The CEEDAR Center's Assessment Practices Within a Multi-Tiered System of Supports tool is one good resource.
- Develop a communication plan for sharing data with families and staff in ways that are clear, culturally responsive, and support student growth.

In our experiences, principals can create a data-informed culture that improves teaching and learning throughout the year by setting a clear vision for schoolwide data use and establishing regular data-review cycles with grade-level teams, special education personnel, and instructional coaches to analyze benchmark data and formative assessments.

Steady efforts to routinely identify students who are struggling and may benefit from Tier 2 or Tier 3 interventions based on multiple forms of data can ensure that struggling learners receive targeted support. These efforts allow principals to make short-term adjustments to professional development and resource allocations that can lead to a better school year.

We hope that recommendations presented here, while not comprehensive, can offer a good starting point for principals seeking to effectively leverage student-achievement data to improve their schools. ■

David E. DeMatthews is the W.K. Kellogg Endowed Professor of Educational Leadership at the University of Texas at Austin. Lebon "Trey" D. James III is an assistant professor and program coordinator for school leadership at Sam Houston State University. Both previously served as school administrators and teachers.



Robert Neubecker for Education Week

Published April 14, 2025

A Strengths-Based Guide to Assessing Student Progress (DOWNLOADABLE)

By Evie Blad & Francis Sheehan

Teachers use rubrics to ensure consistency and clarity in grading student assignments by carefully detailing what a successful project looks like.

But the seeming simplicity of these assessment tools masks complicated questions about how to measure student progress and how to encourage continued learning, said Kevin Perks, the senior director of Quality Schools and Districts at WestEd, a nonpartisan education research organization.

“You have to determine what constructs to measure; how to measure those accurately, with validity; and then how to develop instruments that can be used consistently,” he said. “Assessment literacy is one of the largest knowledge gaps teachers have. It’s an area where they aren’t really trained.”

When designed effectively, rubrics can clarify expectations, minimize subjectivity, and standardize grading criteria across multiple teachers, said Perks, who leads training sessions for teachers about classroom assessment.

There are many schools of thought about how to properly design these grading tools. In this downloadable guide, explore Perks’ suggestions for a simplified, strengths-based rubric centered on a specific learning standard. ■

Additional Resource

See next page for PDF guide



Copyright ©2026 by Editorial Projects in Education, Inc. All rights reserved. No part of this publication shall be reproduced, stored in a retrieval system, or transmitted by any means, electronic or otherwise, without the written permission of the copyright holder.

Readers may make up to 5 print copies of this publication at no cost for personal, non-commercial use, provided that each includes a full citation of the source.

For additional print or electronic copies or to buy in bulk, click [here](#).

Published by Editorial Projects in Education, Inc.
6935 Arlington Road, Suite 100
Bethesda, MD, 20814
Phone: (301) 280-3100
www.edweek.org

Does Your Rubric Make the Grade?

Rethink How You Assess Student Work

When designed effectively, grading rubrics can clarify expectations, minimize subjectivity, and standardize grading criteria across multiple teachers, said Kevin Perks, the senior director of Quality Schools and Districts at WestEd, a nonpartisan education research organization. But traditional rubrics can be confusing and unnecessarily complicated, he said.

Here are Perks' suggestions for a simplified, strengths-based rubric centered on a specific learning standard.

1. Don't add extra language to the learning standard / A typical rubric is a grid with rows of learning expectations on the left side and columns across the top outlining multiple levels of success—such as poor, fair, good, excellent. Teachers fill in each cell on that grid with a description of performance for each expectation at each level. While the aim is clarity, the result can be confusion, Perks said.

"Teachers try to figure out the gradation of language to define each level, and they end up creating new language," he said. "The minute you start adding new language, you are unintentionally adulterating the standard you are trying to teach."

2. Replace terms like "meets expectations" with numbered levels of success / Instead of that traditional grid, Perks recommends a checklist-based scoring guide that breaks down the concepts within a learning standard. Teachers check off demonstrated skills and provide written feedback on how students can improve. Instead of vague labels, number performance levels 1-4, with an index defining each level.

"It's easier for kids to comprehend and it's all really oriented around proficiency," Perks said.

3. Consider co-creating your scoring guide with students / Inviting students to unpack a learning standard at the beginning of a unit can jump-start the learning process and help them understand expectations for assignments, Perks said. Teachers should collaborate with students to identify and define all the concepts within a learning standard and discuss how they might demonstrate those concepts in an assignment. That conversation can shape the rubric or help students understand an existing one.

4. Use examples of strong work / Including examples of successful work alongside scoring criteria can help ensure that different graders apply the same scoring criteria more consistently. For example, the SAT provides "anchors" or example essays for the reviewers to demonstrate various score levels.

Similarly, showing students successful examples of past work can also give them a clearer understanding of the concepts they are learning and how to demonstrate mastery, Perks said. Students may even benefit from reviewing an example that doesn't meet expectations and identifying ways it could be improved.

The Old Way: A Traditional Rubric

	Failing	Basic	Proficient	Exceptional
Concept 1				
Concept 2				
Concept 3				

*Description in each box defines performance for each concept at each level

Try This! A Strengths-Based Rubric

Standard 0.0	☐ Concept 1 ☐ Concept 2 ☐ Concept 3 ☐ Concept 4 ☐ Concept 5 ☐ Concept 6	 Teacher provides feedback on how student can advance learning. For students at level 4, attach the standard for the next grade level.
Scoring index: Level 1- The student work meets some but not all the criteria (less than half of boxes checked). Level 2- The student work meets most but not all the criteria (more than half of boxes checked). Level 3- The student work demonstrates all the knowledge and skills required by the standard. Level 4- In addition to meeting all criteria, the student's work exceeds expectations by demonstrating proficiency in the next grade band and/or by making thoughtful and substantive connections to knowledge and skills in other standards or disciplines.		